



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/31, C07K 14/22, 16/12, C12Q 1/68, A61K 39/095, G01N 33/50		A2	(11) International Publication Number: WO 99/57280
			(43) International Publication Date: 11 November 1999 (11.11.99)
(21) International Application Number: PCT/US99/09346			
(22) International Filing Date: 30 April 1999 (30.04.99)			
(30) Priority Data:			
60/083,758	1 May 1998 (01.05.98)	US	
60/094,869	31 July 1998 (31.07.98)	US	
60/098,994	2 September 1998 (02.09.98)	US	
60/099,062	2 September 1998 (02.09.98)	US	
60/103,749	9 October 1998 (09.10.98)	US	
60/103,794	9 October 1998 (09.10.98)	US	
60/103,796	9 October 1998 (09.10.98)	US	
60/121,528	25 February 1999 (25.02.99)	US	
(71) Applicants (for all designated States except US): CHIRON CORPORATION [US/US]; 4560 Horton Street, Emeryville, CA 94608 (US). THE INSTITUTE FOR GENOMIC RESEARCH [US/US]; 9212 Medical Center Drive, Rockville, MD 20850 (US).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): FRASER, Claire [US/US]; Rockville, MD (US). GALEOTTI, Cesira [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). GRANDI, Guido [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). HICKEY, Erin [US/US]; Gaithersburg, MD			
		(US). MASIGNANI, Vega [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). MORA, Marirosa [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). PETERSEN, Jeremy [US/US]; Arlington, VA (US). PIZZA, Mariagrazia [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RAPPUOLI, Rino [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). RATTI, Giulio [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCALATO, Enzo [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). SCARSELLI, Maria [IT/IT]; Chiron S.p.A., Via Fiorentina, 1, I-53100 Siena (IT). TETTELIN, Herve [US/US]; Gaithersburg, MD (US). VENTER, J., Craig [US/US]; Rockville, MD (US).	
		(74) Agent: HARBIN, Alisa, A.; Chiron Corporation, Intellectual Property - R440, P.O. Box 8097, Emeryville, CA 94662-8097 (US).	
		(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).	
		Published Without international search report and to be republished upon receipt of that report.	
(54) Title: NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS			
(57) Abstract			
The invention provides proteins from Neisseria meningitidis, including the amino acid sequences and the corresponding nucleotide sequences. The proteins are predicted to be useful antigens for vaccines and/or diagnostics.			
<div style="text-align: right;">919 (46 kDa)</div> <div style="text-align: right;">A) PURIFICATION</div> <div style="text-align: right;">M1 919</div>			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakistan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

NEISSERIA MENINGITIDIS ANTIGENS AND COMPOSITIONS

This application is a continuation-in-part of the following U.S. Provisional Patent applications, from each of which priority is claimed, and each of which is incorporated by reference in its entirety: 60/083,758 (filed May 1, 1998); 60/094,869 (filed July 31, 1998); 60/098,994 (filed September 2, 1998); 60/099,062 (filed September 2, 1998); 60/103,749 (filed October 9, 1998); 60/103,794 (filed October 9, 1998); 60/103,796 (filed October 9, 1998); and 60/121,528 (filed February 25, 1999).

This invention relates to antigens from the bacterial species: *Neisseria meningitidis* and *Neisseria gonorrhoeae*.

BACKGROUND

Neisseria meningitidis is a non-motile, gram negative diplococcus human pathogen. It colonizes the pharynx, causing meningitis and, occasionally, septicaemia in the absence of meningitis. It is closely related to *N. gonorrhoea*, although one feature that clearly differentiates meningococcus from gonococcus is the presence of a polysaccharide capsule that is present in all pathogenic meningococci.

N. meningitidis causes both endemic and epidemic disease. In the United States the attack rate is 0.6-1 per 100,000 persons per year, and it can be much greater during outbreaks. (see Lieberman *et al.* (1996) Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. *JAMA* 275(19):1499-1503; Schuchat *et al* (1997) Bacterial Meningitis in the United States in 1995. *N Engl J Med* 337(14):970-976). In developing countries, endemic disease rates are much higher and during epidemics incidence rates can reach 500 cases per 100,000 persons per year. Mortality is extremely high, at 10-20% in the United States, and much higher in developing countries. Following the introduction of the conjugate vaccine against *Haemophilus influenzae*, *N. meningitidis* is the major cause of bacterial meningitis at all ages in the United States (Schuchat *et al* (1997) *supra*).

Based on the organism's capsular polysaccharide, 12 serogroups of *N. meningitidis* have been identified. Group A is the pathogen most often implicated in epidemic disease in sub-Saharan Africa. Serogroups B and C are responsible for the vast majority of cases in the

United States and in most developed countries. Serogroups W135 and Y are responsible for the rest of the cases in the United States and developed countries. The meningococcal vaccine currently in use is a tetravalent polysaccharide vaccine composed of serogroups A, C, Y and W135. Although efficacious in adolescents and adults, it induces a poor immune response and short duration of protection, and cannot be used in infants [eg. Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. This is because polysaccharides are T-cell independent antigens that induce a weak immune response that cannot be boosted by repeated immunization. Following the success of the vaccination against *H.influenzae*, conjugate vaccines against serogroups A and C have been developed and are at the final stage of clinical testing (Zollinger WD "New and Improved Vaccines Against Meningococcal Disease". In: *New Generation Vaccines, supra*, pp. 469-488; Lieberman *et al* (1996) *supra*; Costantino *et al* (1992) Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. *Vaccine* 10:691-698).

Meningococcus B (menB) remains a problem, however. This serotype currently is responsible for approximately 50% of total meningitis in the United States, Europe, and South America. The polysaccharide approach cannot be used because the menB capsular polysaccharide is a polymer of $\alpha(2-8)$ -linked *N*-acetyl neuraminic acid that is also present in mammalian tissue. This results in tolerance to the antigen; indeed, if an immune response were elicited, it would be anti-self, and therefore undesirable. In order to avoid induction of autoimmunity and to induce a protective immune response, the capsular polysaccharide has, for instance, been chemically modified substituting the *N*-acetyl groups with *N*-propionyl groups, leaving the specific antigenicity unaltered (Romero & Outschoorn (1994) Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? *Clin Microbiol Rev* 7(4):559-575).

Alternative approaches to menB vaccines have used complex mixtures of outer membrane proteins (OMPs), containing either the OMPs alone, or OMPs enriched in porins, or deleted of the class 4 OMPs that are believed to induce antibodies that block bactericidal activity. This approach produces vaccines that are not well characterized. They are able to protect against the homologous strain, but are not effective at large where there are many antigenic variants of the outer membrane proteins. To overcome the antigenic variability, multivalent vaccines containing up to nine different porins have been constructed (eg. Poolman JT (1992) Development of a meningococcal vaccine. *Infect. Agents Dis.* 4:13-28).

Additional proteins to be used in outer membrane vaccines have been the opa and opc proteins, but none of these approaches have been able to overcome the antigenic variability (eg. Ala'Aldeen & Borriello (1996) The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. *Vaccine* 14(1):49-53).

A certain amount of sequence data is available for meningococcal and gonococcal genes and proteins (e.g. EP-A-0467714, WO96/29412), but this is by no means complete. Other men B proteins may include those listed in WO 97/28273, WO 96/29412, WO 95/03413, US 5,439,808, and US 5,879,686.

The provision of further sequences could provide an opportunity to identify secreted or surface-exposed proteins that are presumed targets for the immune system and which are not antigenically variable. For instance, some of the identified proteins could be components of efficacious vaccines against meningococcus B, some could be components of vaccines against all meningococcal serotypes, and others could be components of vaccines against all pathogenic *Neisseriae* including *Neisseria meningitidis* or *Neisseria gonorrhoeae*. Those sequences specific to *N. meningitidis* or *N. gonorrhoeae* that are more highly conserved are further preferred sequences.

It is thus an object of the invention is to provide Neisserial DNA sequences which encode proteins that are antigenic or immunogenic.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 illustrates the products of protein expression and purification of the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 2 illustrates the products of protein expression and purification of the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 3 illustrates the products of protein expression and purification of the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 4 illustrates the products of protein expression and purification of the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 5 illustrates the products of protein expression and purification of the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 6 illustrates the products of protein expression and purification of the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 7 illustrates the products of protein expression and purification of the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 8 illustrates the products of protein expression and purification of the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 9 illustrates the products of protein expression and purification of the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 10 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 919 as cloned and expressed in *E. coli*.

Fig. 11 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 279 as cloned and expressed in *E. coli*.

Fig. 12 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 576-1 as cloned and expressed in *E. coli*.

Fig. 13 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 519-1 as cloned and expressed in *E. coli*.

Fig. 14 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 121-1 as cloned and expressed in *E. coli*.

Fig. 15 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 128-1 as cloned and expressed in *E. coli*.

Fig. 16 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 206 as cloned and expressed in *E. coli*.

Fig. 17 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 287 as cloned and expressed in *E. coli*.

Fig. 18 illustrates the hydrophilicity plot, antigenic index and AMPHI regions of the products of protein expression the predicted ORF 406 as cloned and expressed in *E. coli*.

Fig. 19 shows an alignment comparison of amino acid sequences for ORF 225 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 20 shows an alignment comparison of amino acid sequences for ORF 235 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 21 shows an alignment comparison of amino acid sequences for ORF 287 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 22 shows an alignment comparison of amino acid sequences for ORF 519 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

Fig. 23 shows an alignment comparison of amino acid sequences for ORF 919 for several strains of *Neisseria*. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. The Figure demonstrates a high degree of conservation among the various strains, further confirming its utility as an antigen for both vaccines and diagnostics.

THE INVENTION

The invention provides proteins comprising the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples.

It also provides proteins comprising sequences homologous (i.e., those having sequence identity) to the *N. meningitidis* amino acid sequences disclosed in the examples. Depending on the particular sequence, the degree of homology (sequence identity) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). These proteins include mutants and allelic variants of the sequences disclosed in the examples. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence. Identity between proteins is preferably determined by the Smith-Waterman

homology search algorithm as implemented in MPSRCH program (Oxford Molecular) using an affine gap search with parameters: gap penalty 12, gap extension penalty 1.

The invention further provides proteins comprising fragments of the *N. meningitidis* amino acid sequences and *N. gonorrhoeae* amino acid sequences disclosed in the examples. The fragments should comprise at least n consecutive amino acids from the sequences and, depending on the particular sequence, n is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20 or more). Preferably the fragments comprise an epitope from the sequence.

The proteins of the invention can, of course, be prepared by various means (eg. recombinant expression, purification from cell culture, chemical synthesis *etc.*) and in various forms (eg. native, fusions *etc.*). They are preferably prepared in substantially pure or isolated form (*ie.* substantially free from other *N. meningitidis* or *N. gonorrhoeae* host cell proteins)

According to a further aspect, the invention provides antibodies which bind to these proteins. These may be polyclonal or monoclonal and may be produced by any suitable means.

According to a further aspect, the invention provides nucleic acid comprising the *N. meningitidis* nucleotide sequences and *N. gonorrhoeae* nucleotide sequences disclosed in the examples.

According to a further aspect, the invention comprises nucleic acids having sequence identity of greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) to the nucleic acid sequences herein. Sequence identity is determined as above-discussed.

According to a further aspect, the invention comprises nucleic acid that hybridizes to the sequences provided herein. Conditions for hybridization are set forth herein.

Nucleic acid comprising fragments of these sequences are also provided. These should comprise at least n consecutive nucleotides from the *N. meningitidis* sequences or *N. gonorrhoeae* sequences and depending on the particular sequence, n is 10 or more (eg 12, 14, 15, 18, 20, 25, 30, 35, 40 or more).

According to a further aspect, the invention provides nucleic acid encoding the proteins and protein fragments of the invention.

It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (eg. for antisense or probing purposes).

Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, in part or in whole, from genomic or cDNA libraries, from the

organism itself *etc.*) and can take various forms (*eg.* single stranded, double stranded, vectors, probes *etc.*).

In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also protein nucleic acids (PNA) *etc.*

According to a further aspect, the invention provides vectors comprising nucleotide sequences of the invention (*eg.* expression vectors) and host cells transformed with such vectors.

According to a further aspect, the invention provides compositions comprising protein, antibody, and/or nucleic acid according to the invention. These compositions may be suitable as vaccines, for instance, or as diagnostic reagents or as immunogenic compositions.

The invention also provides nucleic acid, protein, or antibody according to the invention for use as medicaments (*eg.* as vaccines) or as diagnostic reagents. It also provides the use of nucleic acid, protein, or antibody according to the invention in the manufacture of (I) a medicament for treating or preventing infection due to Neisserial bacteria (ii) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria or (iii) for raising antibodies. Said Neisserial bacteria may be any species or strain (such as *N. gonorrhoeae*) but are preferably *N. meningitidis*, especially strain B or strain C.

The invention also provides a method of treating a patient, comprising administering to the patient a therapeutically effective amount of nucleic acid, protein, and/or antibody according to the invention.

According to further aspects, the invention provides various processes.

A process for producing proteins of the invention is provided, comprising the step of culturing a host cell according to the invention under conditions which induce protein expression.

A process for detecting polynucleotides of the invention is provided, comprising the steps of: (a) contacting a nucleic probe according to the invention with a biological sample under hybridizing conditions to form duplexes; and (b) detecting said duplexes.

A process for detecting proteins of the invention is provided, comprising the steps of: (a) contacting an antibody according to the invention with a biological sample under conditions suitable for the formation of an antibody-antigen complexes; and (b) detecting said complexes.

Having now generally described the invention, the same will be more readily understood through reference to the following examples which are provided by way of illustration, and are not intended to be limiting of the present invention, unless specified.

Methodology - Summary of standard procedures and techniques.

General

This invention provides *Neisseria meningitidis* menB nucleotide sequences, amino acid sequences encoded therein. With these disclosed sequences, nucleic acid probe assays and expression cassettes and vectors can be produced. The expression vectors can be transformed into host cells to produce proteins. The purified or isolated polypeptides (which may also be chemically synthesized) can be used to produce antibodies to detect menB proteins. Also, the host cells or extracts can be utilized for biological assays to isolate agonists or antagonists. In addition, with these sequences one can search to identify open reading frames and identify amino acid sequences. The proteins may also be used in immunogenic compositions, antigenic compositions and as vaccine components.

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of molecular biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature e.g., Sambrook *Molecular Cloning; A Laboratory Manual, Second Edition* (1989); *DNA Cloning, Volumes I and ii* (D.N Glover ed. 1985); *Oligonucleotide Synthesis* (M.J. Gait ed, 1984); *Nucleic Acid Hybridization* (B.D. Hames & S.J. Higgins eds. 1984); *Transcription and Translation* (B.D. Hames & S.J. Higgins eds. 1984); *Animal Cell Culture* (R.I. Freshney ed. 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); the *Methods in Enzymology* series (Academic Press, Inc.), especially volumes 154 & 155; *Gene Transfer Vectors for Mammalian Cells* (J.H. Miller and M.P. Calos eds. 1987, Cold Spring Harbor Laboratory); Mayer and Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes, (1987) *Protein Purification: Principles and Practice*, Second Edition (Springer-Verlag, N.Y.), and *Handbook of Experimental Immunology, Volumes I-IV* (D.M. Weir and C.C. Blackwell eds 1986).

Standard abbreviations for nucleotides and amino acids are used in this specification.

All publications, patents, and patent applications cited herein are incorporated in full by reference.

Expression systems

The *Neisseria* menB nucleotide sequences can be expressed in a variety of different expression systems; for example those used with mammalian cells, plant cells, baculoviruses, bacteria, and yeast.

i. Mammalian Systems

Mammalian expression systems are known in the art. A mammalian promoter is any DNA sequence capable of binding mammalian RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiating region, which is usually placed proximal to the 5' end of the coding sequence, and a TATA box, usually located 25-30 base pairs (bp) upstream of the transcription initiation site. The TATA box is thought to direct RNA polymerase II to begin RNA synthesis at the correct site. A mammalian promoter will also contain an upstream promoter element, usually located within 100 to 200 bp upstream of the TATA box. An upstream promoter element determines the rate at which transcription is initiated and can act in either orientation (Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells." In *Molecular Cloning: A Laboratory Manual*, 2nd ed.).

Mammalian viral genes are often highly expressed and have a broad host range; therefore sequences encoding mammalian viral genes provide particularly useful promoter sequences. Examples include the SV40 early promoter, mouse mammary tumor virus LTR promoter, adenovirus major late promoter (Ad MLP), and herpes simplex virus promoter. In addition, sequences derived from non-viral genes, such as the murine metallothionein gene, also provide useful promoter sequences. Expression may be either constitutive or regulated (inducible). Depending on the promoter selected, many promoters may be inducible using known substrates, such as the use of the mouse mammary tumor virus (MMTV) promoter with the glucocorticoid responsive element (GRE) that is induced by glucocorticoid in hormone-responsive transformed cells (see for example, U.S. Patent 5,783,681).

The presence of an enhancer element (enhancer), combined with the promoter elements described above, will usually increase expression levels. An enhancer is a

regulatory DNA sequence that can stimulate transcription up to 1000-fold when linked to homologous or heterologous promoters, with synthesis beginning at the normal RNA start site. Enhancers are also active when they are placed upstream or downstream from the transcription initiation site, in either normal or flipped orientation, or at a distance of more than 1000 nucleotides from the promoter (Maniatis et al. (1987) *Science* 236:1237; Alberts et al. (1989) *Molecular Biology of the Cell*, 2nd ed.). Enhancer elements derived from viruses may be particularly useful, because they usually have a broader host range. Examples include the SV40 early gene enhancer (Dijkema et al (1985) *EMBO J.* 4:761) and the enhancer/promoters derived from the long terminal repeat (LTR) of the Rous Sarcoma Virus (Gorman et al. (1982b) *Proc. Natl. Acad. Sci.* 79:6777) and from human cytomegalovirus (Boshart et al. (1985) *Cell* 41:521). Additionally, some enhancers are regulatable and become active only in the presence of an inducer, such as a hormone or metal ion (Sassone-Corsi and Borelli (1986) *Trends Genet.* 2:215; Maniatis et al. (1987) *Science* 236:1237).

A DNA molecule may be expressed intracellularly in mammalian cells. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in mammalian cells. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The adenovirus tripartite leader is an example of a leader sequence that provides for secretion of a foreign protein in mammalian cells.

Usually, transcription termination and polyadenylation sequences recognized by mammalian cells are regulatory regions located 3' to the translation stop codon and thus, together with the promoter elements, flank the coding sequence. The 3' terminus of the mature mRNA is formed by site-specific post-transcriptional cleavage and polyadenylation (Birnstiel et al. (1985) *Cell* 41:349; Proudfoot and Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA. In *Transcription and splicing* (ed. B.D. Hames and D.M.

Glover); Proudfoot (1989) *Trends Biochem. Sci.* 14:105). These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator/polyadenylation signals include those derived from SV40 (Sambrook et al (1989) "Expression of cloned genes in cultured mammalian cells." In *Molecular Cloning: A Laboratory Manual*).

Usually, the above described components, comprising a promoter, polyadenylation signal, and transcription termination sequence are put together into expression constructs. Enhancers, introns with functional splice donor and acceptor sites, and leader sequences may also be included in an expression construct, if desired. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as mammalian cells or bacteria. Mammalian replication systems include those derived from animal viruses, which require trans-acting factors to replicate. For example, plasmids containing the replication systems of papovaviruses, such as SV40 (Gluzman (1981) *Cell* 23:175) or polyomavirus, replicate to extremely high copy number in the presence of the appropriate viral T antigen. Additional examples of mammalian replicons include those derived from bovine papillomavirus and Epstein-Barr virus. Additionally, the replicon may have two replication systems, thus allowing it to be maintained, for example, in mammalian cells for expression and in a prokaryotic host for cloning and amplification. Examples of such mammalian-bacteria shuttle vectors include pMT2 (Kaufman et al. (1989) *Mol. Cell. Biol.* 9:946) and pHEBO (Shimizu et al. (1986) *Mol. Cell. Biol.* 6:1074).

The transformation procedure used depends upon the host to be transformed. Methods for introduction of heterologous polynucleotides into mammalian cells are known in the art and include dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei.

Mammalian cell lines available as hosts for expression are known in the art and include many immortalized cell lines available from the American Type Culture Collection (ATCC), including but not limited to, Chinese hamster ovary (CHO) cells, HeLa cells, baby hamster kidney (BHK) cells, monkey kidney cells (COS), human hepatocellular carcinoma cells (e.g., Hep G2), and a number of other cell lines.

ii. Plant Cellular Expression Systems

There are many plant cell culture and whole plant genetic expression systems known in the art. Exemplary plant cellular genetic expression systems include those described in patents, such as: U.S. 5,693,506; US 5,659,122; and US 5,608,143. Additional examples of genetic expression in plant cell culture has been described by Zenk, *Phytochemistry* 30:3861-3863 (1991). Descriptions of plant protein signal peptides may be found in addition to the references described above in Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, *J. Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., *Nucleic Acids Research* 15:2515-2535 (1987); Wirsal et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). A description of the regulation of plant gene expression by the phytohormone, gibberellic acid and secreted enzymes induced by gibberellic acid can be found in R.L. Jones and J. MacMillin, Gibberellins: in: *Advanced Plant Physiology*, Malcolm B. Wilkins, ed., 1984 Pitman Publishing Limited, London, pp. 21-52. References that describe other metabolically-regulated genes: Sheen, *Plant Cell*, 2:1027-1038(1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel and Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987)

Typically, using techniques known in the art, a desired polynucleotide sequence is inserted into an expression cassette comprising genetic regulatory elements designed for operation in plants. The expression cassette is inserted into a desired expression vector with companion sequences upstream and downstream from the expression cassette suitable for expression in a plant host. The companion sequences will be of plasmid or viral origin and provide necessary characteristics to the vector to permit the vectors to move DNA from an original cloning host, such as bacteria, to the desired plant host. The basic bacterial/plant vector construct will preferably provide a broad host range prokaryote replication origin; a prokaryote selectable marker; and, for *Agrobacterium* transformations, T DNA sequences for *Agrobacterium*-mediated transfer to plant chromosomes. Where the heterologous gene is not readily amenable to detection, the construct will preferably also have a selectable marker gene suitable for determining if a plant cell has been transformed. A general review of suitable markers, for example for the members of the grass family, is found in Wilmink and Dons, 1993, *Plant Mol. Biol. Repr.*, 11(2):165-185.

Sequences suitable for permitting integration of the heterologous sequence into the plant genome are also recommended. These might include transposon sequences and the like for homologous recombination as well as Ti sequences which permit random insertion of a heterologous expression cassette into a plant genome. Suitable prokaryote selectable markers include resistance toward antibiotics such as ampicillin or tetracycline. Other DNA sequences encoding additional functions may also be present in the vector, as is known in the art.

The nucleic acid molecules of the subject invention may be included into an expression cassette for expression of the protein(s) of interest. Usually, there will be only one expression cassette, although two or more are feasible. The recombinant expression cassette will contain in addition to the heterologous protein encoding sequence the following elements, a promoter region, plant 5' untranslated sequences, initiation codon depending upon whether or not the structural gene comes equipped with one, and a transcription and translation termination sequence. Unique restriction enzyme sites at the 5' and 3' ends of the cassette allow for easy insertion into a pre-existing vector.

A heterologous coding sequence may be for any protein relating to the present invention. The sequence encoding the protein of interest will encode a signal peptide which allows processing and translocation of the protein, as appropriate, and will usually lack any sequence which might result in the binding of the desired protein of the invention to a membrane. Since, for the most part, the transcriptional initiation region will be for a gene which is expressed and translocated during germination, by employing the signal peptide which provides for translocation, one may also provide for translocation of the protein of interest. In this way, the protein(s) of interest will be translocated from the cells in which they are expressed and may be efficiently harvested. Typically secretion in seeds are across the aleurone or scutellar epithelium layer into the endosperm of the seed. While it is not required that the protein be secreted from the cells in which the protein is produced, this facilitates the isolation and purification of the recombinant protein.

Since the ultimate expression of the desired gene product will be in a eucaryotic cell it is desirable to determine whether any portion of the cloned gene contains sequences which will be processed out as introns by the host's splicosome machinery. If so, site-directed mutagenesis of the "intron" region may be conducted to prevent losing a portion of the genetic message as a false intron code, Reed and Maniatis, *Cell* 41:95-105, 1985.

The vector can be microinjected directly into plant cells by use of micropipettes to mechanically transfer the recombinant DNA. Crossway, *Mol. Gen. Genet*, 202:179-185, 1985. The genetic material may also be transferred into the plant cell by using polyethylene glycol, Krens, et al., *Nature*, 296, 72-74, 1982. Another method of introduction of nucleic acid segments is high velocity ballistic penetration by small particles with the nucleic acid either within the matrix of small beads or particles, or on the surface, Klein, et al., *Nature*, 327, 70-73, 1987 and Knudsen and Muller, 1991, *Planta*, 185:330-336 teaching particle bombardment of barley endosperm to create transgenic barley. Yet another method of introduction would be fusion of protoplasts with other entities, either minicells, cells, lysosomes or other fusible lipid-surfaced bodies, Fraley, et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

The vector may also be introduced into the plant cells by electroporation. (Fromm et al., *Proc. Natl. Acad. Sci. USA* 82:5824, 1985). In this technique, plant protoplasts are electroporated in the presence of plasmids containing the gene construct. Electrical impulses of high field strength reversibly permeabilize biomembranes allowing the introduction of the plasmids. Electroporated plant protoplasts reform the cell wall, divide, and form plant callus.

All plants from which protoplasts can be isolated and cultured to give whole regenerated plants can be transformed by the present invention so that whole plants are recovered which contain the transferred gene. It is known that practically all plants can be regenerated from cultured cells or tissues, including but not limited to all major species of sugarcane, sugar beet, cotton, fruit and other trees, legumes and vegetables. Some suitable plants include, for example, species from the genera *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum*, *Sorghum*, and *Datura*.

Means for regeneration vary from species to species of plants, but generally a suspension of transformed protoplasts containing copies of the heterologous gene is first provided. Callus tissue is formed and shoots may be induced from callus and subsequently rooted. Alternatively, embryo formation can be induced from the protoplast suspension.

These embryos germinate as natural embryos to form plants. The culture media will generally contain various amino acids and hormones, such as auxin and cytokinins. It is also advantageous to add glutamic acid and proline to the medium, especially for such species as corn and alfalfa. Shoots and roots normally develop simultaneously. Efficient regeneration will depend on the medium, on the genotype, and on the history of the culture. If these three variables are controlled, then regeneration is fully reproducible and repeatable.

In some plant cell culture systems, the desired protein of the invention may be excreted or alternatively, the protein may be extracted from the whole plant. Where the desired protein of the invention is secreted into the medium, it may be collected. Alternatively, the embryos and embryoless-half seeds or other plant tissue may be mechanically disrupted to release any secreted protein between cells and tissues. The mixture may be suspended in a buffer solution to retrieve soluble proteins. Conventional protein isolation and purification methods will be then used to purify the recombinant protein. Parameters of time, temperature pH, oxygen, and volumes will be adjusted through routine methods to optimize expression and recovery of heterologous protein.

iii. Baculovirus Systems

The polynucleotide encoding the protein can also be inserted into a suitable insect expression vector, and is operably linked to the control elements within that vector. Vector construction employs techniques which are known in the art. Generally, the components of the expression system include a transfer vector, usually a bacterial plasmid, which contains both a fragment of the baculovirus genome, and a convenient restriction site for insertion of the heterologous gene or genes to be expressed; a wild type baculovirus with a sequence homologous to the baculovirus-specific fragment in the transfer vector (this allows for the homologous recombination of the heterologous gene in to the baculovirus genome); and appropriate insect host cells and growth media.

After inserting the DNA sequence encoding the protein into the transfer vector, the vector and the wild type viral genome are transfected into an insect host cell where the vector and viral genome are allowed to recombine. The packaged recombinant virus is expressed and recombinant plaques are identified and purified. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *inter alia*, Invitrogen, San Diego CA ("MaxBac" kit). These techniques are generally known to

those skilled in the art and fully described in Summers and Smith, *Texas Agricultural Experiment Station Bulletin No. 1555* (1987) (hereinafter "Summers and Smith").

Prior to inserting the DNA sequence encoding the protein into the baculovirus genome, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are usually assembled into an intermediate transplacement construct (transfer vector). This construct may contain a single gene and operably linked regulatory elements; multiple genes, each with its own set of operably linked regulatory elements; or multiple genes, regulated by the same set of regulatory elements. Intermediate transplacement constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as a bacterium. The replicon will have a replication system, thus allowing it to be maintained in a suitable host for cloning and amplification.

Currently, the most commonly used transfer vector for introducing foreign genes into AcNPV is pAc373. Many other vectors, known to those of skill in the art, have also been designed. These include, for example, pVL985 (which alters the polyhedrin start codon from ATG to ATT, and which introduces a BamHI cloning site 32 basepairs downstream from the ATT; see Luckow and Summers, *Virology* (1989) 17:31.

The plasmid usually also contains the polyhedrin polyadenylation signal (Miller et al. (1988) *Ann. Rev. Microbiol.*, 42:177) and a prokaryotic ampicillin-resistance (*amp*) gene and origin of replication for selection and propagation in *E. coli*.

Baculovirus transfer vectors usually contain a baculovirus promoter. A baculovirus promoter is any DNA sequence capable of binding a baculovirus RNA polymerase and initiating the downstream (5' to 3') transcription of a coding sequence (e.g., structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A baculovirus transfer vector may also have a second domain called an enhancer, which, if present, is usually distal to the structural gene. Expression may be either regulated or constitutive.

Structural genes, abundantly transcribed at late times in a viral infection cycle, provide particularly useful promoter sequences. Examples include sequences derived from the gene encoding the viral polyhedron protein, Friesen et al., (1986) "The Regulation of Baculovirus Gene Expression," in: *The Molecular Biology of Baculoviruses* (ed. Walter Doerfler); EPO

Publ. Nos. 127 839 and 155 476; and the gene encoding the p10 protein, Vlcek et al., (1988), *J. Gen. Virol.* 69:765.

DNA encoding suitable signal sequences can be derived from genes for secreted insect or baculovirus proteins, such as the baculovirus polyhedrin gene (Carbonell et al. (1988) *Gene*, 73:409). Alternatively, since the signals for mammalian cell posttranslational modifications (such as signal peptide cleavage, proteolytic cleavage, and phosphorylation) appear to be recognized by insect cells, and the signals required for secretion and nuclear accumulation also appear to be conserved between the invertebrate cells and vertebrate cells, leaders of non-insect origin, such as those derived from genes encoding human (alpha) α -interferon, Maeda et al., (1985), *Nature* 315:592; human gastrin-releasing peptide, Lebacqz-Verheyden et al., (1988), *Molec. Cell. Biol.* 8:3129; human IL-2, Smith et al., (1985) *Proc. Nat'l Acad. Sci. USA*, 82:8404; mouse IL-3, (Miyajima et al., (1987) *Gene* 58:273; and human glucocerebrosidase, Martin et al. (1988) *DNA*, 7:99, can also be used to provide for secretion in insects.

A recombinant polypeptide or polyprotein may be expressed intracellularly or, if it is expressed with the proper regulatory sequences, it can be secreted. Good intracellular expression of nonfused foreign proteins usually requires heterologous genes that ideally have a short leader sequence containing suitable translation initiation signals preceding an ATG start signal. If desired, methionine at the N-terminus may be cleaved from the mature protein by *in vitro* incubation with cyanogen bromide.

Alternatively, recombinant polyproteins or proteins which are not naturally secreted can be secreted from the insect cell by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provides for secretion of the foreign protein in insects. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the translocation of the protein into the endoplasmic reticulum.

After insertion of the DNA sequence and/or the gene encoding the expression product precursor of the protein, an insect cell host is co-transformed with the heterologous DNA of the transfer vector and the genomic DNA of wild type baculovirus -- usually by co-transfection. The promoter and transcription termination sequence of the construct will usually comprise a 2-5kb section of the baculovirus genome. Methods for introducing heterologous DNA into the desired site in the baculovirus virus are known in the art. (See

Summers and Smith *supra*; Ju et al. (1987); Smith et al., *Mol. Cell. Biol.* (1983) 3:2156; and Luckow and Summers (1989)). For example, the insertion can be into a gene such as the polyhedrin gene, by homologous double crossover recombination; insertion can also be into a restriction enzyme site engineered into the desired baculovirus gene. Miller et al., (1989), *Bioessays* 4:91. The DNA sequence, when cloned in place of the polyhedrin gene in the expression vector, is flanked both 5' and 3' by polyhedrin-specific sequences and is positioned downstream of the polyhedrin promoter.

The newly formed baculovirus expression vector is subsequently packaged into an infectious recombinant baculovirus. Homologous recombination occurs at low frequency (between about 1% and about 5%); thus, the majority of the virus produced after cotransfection is still wild-type virus. Therefore, a method is necessary to identify recombinant viruses. An advantage of the expression system is a visual screen allowing recombinant viruses to be distinguished. The polyhedrin protein, which is produced by the native virus, is produced at very high levels in the nuclei of infected cells at late times after viral infection. Accumulated polyhedrin protein forms occlusion bodies that also contain embedded particles. These occlusion bodies, up to 15 μm in size, are highly refractile, giving them a bright shiny appearance that is readily visualized under the light microscope. Cells infected with recombinant viruses lack occlusion bodies. To distinguish recombinant virus from wild-type virus, the transfection supernatant is plaqued onto a monolayer of insect cells by techniques known to those skilled in the art. Namely, the plaques are screened under the light microscope for the presence (indicative of wild-type virus) or absence (indicative of recombinant virus) of occlusion bodies. *Current Protocols in Microbiology* Vol. 2 (Ausubel et al. eds) at 16.8 (Supp. 10, 1990); Summers and Smith, *supra*; Miller et al. (1989).

Recombinant baculovirus expression vectors have been developed for infection into several insect cells. For example, recombinant baculoviruses have been developed for, *inter alia*: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda*, and *Trichoplusia ni* (PCT Pub. No. WO 89/046699; Carbonell et al., (1985) *J. Virol.* 56:153; Wright (1986) *Nature* 321:718; Smith et al., (1983) *Mol. Cell. Biol.* 3:2156; and see generally, Fraser, et al. (1989) *In Vitro Cell. Dev. Biol.* 25:225).

Cells and cell culture media are commercially available for both direct and fusion expression of heterologous polypeptides in a baculovirus/expression system; cell culture technology is generally known to those skilled in the art. See, e.g., Summers and Smith *supra*.

The modified insect cells may then be grown in an appropriate nutrient medium, which allows for stable maintenance of the plasmid(s) present in the modified insect host. Where the expression product gene is under inducible control, the host may be grown to high density, and expression induced. Alternatively, where expression is constitutive, the product will be continuously expressed into the medium and the nutrient medium must be continuously circulated, while removing the product of interest and augmenting depleted nutrients. The product may be purified by such techniques as chromatography, e.g., HPLC, affinity chromatography, ion exchange chromatography, etc.; electrophoresis; density gradient centrifugation; solvent extraction, or the like. As appropriate, the product may be further purified, as required, so as to remove substantially any insect proteins which are also secreted in the medium or result from lysis of insect cells, so as to provide a product which is at least substantially free of host debris, e.g., proteins, lipids and polysaccharides.

In order to obtain protein expression, recombinant host cells derived from the transformants are incubated under conditions which allow expression of the recombinant protein encoding sequence. These conditions will vary, dependent upon the host cell selected. However, the conditions are readily ascertainable to those of ordinary skill in the art, based upon what is known in the art.

iv. Bacterial Systems

Bacterial expression techniques are known in the art. A bacterial promoter is any DNA sequence capable of binding bacterial RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site and a transcription initiation site. A bacterial promoter may also have a second domain called an operator, that may overlap an adjacent RNA polymerase binding site at which RNA synthesis begins. The operator permits negative regulated (inducible) transcription, as a gene repressor protein may bind the operator and thereby inhibit transcription of a specific gene. Constitutive expression may occur in the absence of negative regulatory elements, such as the operator. In addition, positive regulation may be achieved by a gene activator protein binding sequence, which, if present is usually proximal (5') to the RNA polymerase binding sequence. An example of a gene activator protein is the catabolite activator protein (CAP), which helps

initiate transcription of the lac operon in *Escherichia coli* (*E. coli*) (Raibaud *et al.* (1984) *Annu. Rev. Genet.* 18:173). Regulated expression may therefore be either positive or negative, thereby either enhancing or reducing transcription.

Sequences encoding metabolic pathway enzymes provide particularly useful promoter sequences. Examples include promoter sequences derived from sugar metabolizing enzymes, such as galactose, lactose (*lac*) (Chang *et al.* (1977) *Nature* 198:1056), and maltose. Additional examples include promoter sequences derived from biosynthetic enzymes such as tryptophan (*trp*) (Goeddel *et al.* (1980) *Nuc. Acids Res.* 8:4057; Yelverton *et al.* (1981) *Nucl. Acids Res.* 9:731; U.S. Patent 4,738,921; EPO Publ. Nos. 036 776 and 121 775). The beta-lactamase (*bla*) promoter system (Weissmann (1981) "The cloning of interferon and other mistakes." In *Interferon 3* (ed. I. Gresser)), bacteriophage lambda PL (Shimatake *et al.* (1981) *Nature* 292:128) and T5 (U.S. Patent 4,689,406) promoter systems also provide useful promoter sequences.

In addition, synthetic promoters which do not occur in nature also function as bacterial promoters. For example, transcription activation sequences of one bacterial or bacteriophage promoter may be joined with the operon sequences of another bacterial or bacteriophage promoter, creating a synthetic hybrid promoter (U.S. Patent 4,551,433). For example, the *tac* promoter is a hybrid *trp-lac* promoter comprised of both *trp* promoter and *lac* operon sequences that is regulated by the *lac* repressor (Amann *et al.* (1983) *Gene* 25:167; de Boer *et al.* (1983) *Proc. Natl. Acad. Sci.* 80:21). Furthermore, a bacterial promoter can include naturally occurring promoters of non-bacterial origin that have the ability to bind bacterial RNA polymerase and initiate transcription. A naturally occurring promoter of non-bacterial origin can also be coupled with a compatible RNA polymerase to produce high levels of expression of some genes in prokaryotes. The bacteriophage T7 RNA polymerase/promoter system is an example of a coupled promoter system (Studier *et al.* (1986) *J. Mol. Biol.* 189:113; Tabor *et al.* (1985) *Proc Natl. Acad. Sci.* 82:1074). In addition, a hybrid promoter can also be comprised of a bacteriophage promoter and an *E. coli* operator region (EPO Publ. No. 267 851).

In addition to a functioning promoter sequence, an efficient ribosome binding site is also useful for the expression of foreign genes in prokaryotes. In *E. coli*, the ribosome binding site is called the Shine-Dalgarno (SD) sequence and includes an initiation codon (ATG) and a sequence 3-9 nucleotides in length located 3-11 nucleotides upstream of the initiation codon

(Shine *et al.* (1975) *Nature* 254:34). The SD sequence is thought to promote binding of mRNA to the ribosome by the pairing of bases between the SD sequence and the 3' end of *E. coli* 16S rRNA (Steitz *et al.* (1979) "Genetic signals and nucleotide sequences in messenger RNA." In *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberg)). To express eukaryotic genes and prokaryotic genes with weak ribosome-binding site, it is often necessary to optimize the distance between the SD sequence and the ATG of the eukaryotic gene (Sambrook *et al.* (1989) "Expression of cloned genes in *Escherichia coli*." In *Molecular Cloning: A Laboratory Manual*).

A DNA molecule may be expressed intracellularly. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide or by either *in vivo* or *in vitro* incubation with a bacterial methionine N-terminal peptidase (EPO Publ. No. 219 237).

Fusion proteins provide an alternative to direct expression. Usually, a DNA sequence encoding the N-terminal portion of an endogenous bacterial protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the bacteriophage lambda cell gene can be linked at the 5' terminus of a foreign gene and expressed in bacteria. The resulting fusion protein preferably retains a site for a processing enzyme (factor Xa) to cleave the bacteriophage protein from the foreign gene (Nagai *et al.* (1984) *Nature* 309:810). Fusion proteins can also be made with sequences from the *lacZ* (Jia *et al.* (1987) *Gene* 60:197), *trpE* (Allen *et al.* (1987) *J. Biotechnol.* 5:93; Makoff *et al.* (1989) *J. Gen. Microbiol.* 135:11), and *Chey* (EPO Publ. No. 324 647) genes. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin specific processing-protease) to cleave the ubiquitin from the foreign protein. Through this method, native foreign protein can be isolated (Miller *et al.* (1989) *Bio/Technology* 7:698).

Alternatively, foreign proteins can also be secreted from the cell by creating chimeric DNA molecules that encode a fusion protein comprised of a signal peptide sequence fragment that provides for secretion of the foreign protein in bacteria (U.S. Patent 4,336,336). The

signal sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell. The protein is either secreted into the growth media (gram-positive bacteria) or into the periplasmic space, located between the inner and outer membrane of the cell (gram-negative bacteria). Preferably there are processing sites, which can be cleaved either *in vivo* or *in vitro* encoded between the signal peptide fragment and the foreign gene.

DNA encoding suitable signal sequences can be derived from genes for secreted bacterial proteins, such as the *E. coli* outer membrane protein gene (*ompA*) (Masui *et al.* (1983), in: *Experimental Manipulation of Gene Expression*; Ghayeb *et al.* (1984) *EMBO J.* 3:2437) and the *E. coli* alkaline phosphatase signal sequence (*phoA*) (Oka *et al.* (1985) *Proc. Natl. Acad. Sci.* 82:7212). As an additional example, the signal sequence of the alpha-amylase gene from various *Bacillus* strains can be used to secrete heterologous proteins from *B. subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. No. 244 042).

Usually, transcription termination sequences recognized by bacteria are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Transcription termination sequences frequently include DNA sequences of about 50 nucleotides capable of forming stem loop structures that aid in terminating transcription. Examples include transcription termination sequences derived from genes with strong promoters, such as the *trp* gene in *E. coli* as well as other biosynthetic genes.

Usually, the above described components, comprising a promoter, signal sequence (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as bacteria. The replicon will have a replication system, thus allowing it to be maintained in a prokaryotic host either for expression or for cloning and amplification. In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably contain at least about 10, and more preferably at least about 20 plasmids. Either a high or low copy number

vector may be selected, depending upon the effect of the vector and the foreign protein on the host.

Alternatively, the expression constructs can be integrated into the bacterial genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to the bacterial chromosome that allows the vector to integrate. Integrations appear to result from recombinations between homologous DNA in the vector and the bacterial chromosome. For example, integrating vectors constructed with DNA from various *Bacillus* strains integrate into the *Bacillus* chromosome (EPO Publ. No. 127 328). Integrating vectors may also be comprised of bacteriophage or transposon sequences.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of bacterial strains that have been transformed. Selectable markers can be expressed in the bacterial host and may include genes which render bacteria resistant to drugs such as ampicillin, chloramphenicol, erythromycin, kanamycin (neomycin), and tetracycline (Davies *et al.* (1978) *Annu. Rev. Microbiol.* 32:469). Selectable markers may also include biosynthetic genes, such as those in the histidine, tryptophan, and leucine biosynthetic pathways.

Alternatively, some of the above described components can be put together in transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extra-chromosomal replicons or integrating vectors, have been developed for transformation into many bacteria. For example, expression vectors have been developed for, *inter alia*, the following bacteria: *Bacillus subtilis* (Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541), *Escherichia coli* (Shimatake *et al.* (1981) *Nature* 292:128; Amann *et al.* (1985) *Gene* 40:183; Studier *et al.* (1986) *J. Mol. Biol.* 189:113; EPO Publ. Nos. 036 776, 136 829 and 136 907), *Streptococcus cremoris* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655); *Streptococcus lividans* (Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655), *Streptomyces lividans* (U.S. Patent 4,745,056).

Methods of introducing exogenous DNA into bacterial hosts are well-known in the art, and usually include either the transformation of bacteria treated with CaCl_2 or other agents, such as divalent cations and DMSO. DNA can also be introduced into bacterial cells by

electroporation. Transformation procedures usually vary with the bacterial species to be transformed. (See e.g., use of *Bacillus*: Masson *et al.* (1989) *FEMS Microbiol. Lett.* 60:273; Palva *et al.* (1982) *Proc. Natl. Acad. Sci. USA* 79:5582; EPO Publ. Nos. 036 259 and 063 953; PCT Publ. No. WO 84/04541; use of *Campylobacter*: Miller *et al.* (1988) *Proc. Natl. Acad. Sci.* 85:856; and Wang *et al.* (1990) *J. Bacteriol.* 172:949; use of *Escherichia coli*: Cohen *et al.* (1973) *Proc. Natl. Acad. Sci.* 69:2110; Dower *et al.* (1988) *Nucleic Acids Res.* 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. In *Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering* (eds. H.W. Boyer and S. Nicosia); Mandel *et al.* (1970) *J. Mol. Biol.* 53:159; Taketo (1988) *Biochim. Biophys. Acta* 949:318; use of *Lactobacillus*: Chassy *et al.* (1987) *FEMS Microbiol. Lett.* 44:173; use of *Pseudomonas*: Fiedler *et al.* (1988) *Anal. Biochem* 170:38; use of *Staphylococcus*: Augustin *et al.* (1990) *FEMS Microbiol. Lett.* 66:203; use of *Streptococcus*: Barany *et al.* (1980) *J. Bacteriol.* 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation, in: *Streptococcal Genetics* (ed. J. Ferretti and R. Curtiss III); Perry *et al.* (1981) *Infect. Immun.* 32:1295; Powell *et al.* (1988) *Appl. Environ. Microbiol.* 54:655; Somkuti *et al.* (1987) *Proc. 4th Evr. Cong. Biotechnology* 1:412.

v. Yeast Expression

Yeast expression systems are also known to one of ordinary skill in the art. A yeast promoter is any DNA sequence capable of binding yeast RNA polymerase and initiating the downstream (3') transcription of a coding sequence (e.g. structural gene) into mRNA. A promoter will have a transcription initiation region which is usually placed proximal to the 5' end of the coding sequence. This transcription initiation region usually includes an RNA polymerase binding site (the "TATA Box") and a transcription initiation site. A yeast promoter may also have a second domain called an upstream activator sequence (UAS), which, if present, is usually distal to the structural gene. The UAS permits regulated (inducible) expression. Constitutive expression occurs in the absence of a UAS. Regulated expression may be either positive or negative, thereby either enhancing or reducing transcription.

Yeast is a fermenting organism with an active metabolic pathway, therefore sequences encoding enzymes in the metabolic pathway provide particularly useful promoter sequences.

Examples include alcohol dehydrogenase (ADH) (EPO Publ. No. 284 044), enolase, glucokinase, glucose-6-phosphate isomerase, glyceraldehyde-3-phosphate-dehydrogenase (GAP or GAPDH), hexokinase, phosphofructokinase, 3-phosphoglycerate mutase, and pyruvate kinase (PyK) (EPO Publ. No. 329 203). The yeast *PHO5* gene, encoding acid phosphatase, also provides useful promoter sequences (Myanohara *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:1).

In addition, synthetic promoters which do not occur in nature also function as yeast promoters. For example, UAS sequences of one yeast promoter may be joined with the transcription activation region of another yeast promoter, creating a synthetic hybrid promoter. Examples of such hybrid promoters include the ADH regulatory sequence linked to the GAP transcription activation region (U.S. Patent Nos. 4,876,197 and 4,880,734). Other examples of hybrid promoters include promoters which consist of the regulatory sequences of either the *ADH2*, *GAL4*, *GAL10*, OR *PHO5* genes, combined with the transcriptional activation region of a glycolytic enzyme gene such as GAP or PyK (EPO Publ. No. 164 556). Furthermore, a yeast promoter can include naturally occurring promoters of non-yeast origin that have the ability to bind yeast RNA polymerase and initiate transcription. Examples of such promoters include, *inter alia*, (Cohen *et al.* (1980) *Proc. Natl. Acad. Sci. USA* 77:1078; Henikoff *et al.* (1981) *Nature* 283:835; Hollenberg *et al.* (1981) *Curr. Topics Microbiol. Immunol.* 96:119; Hollenberg *et al.* (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast *Saccharomyces cerevisiae*," in: *Plasmids of Medical, Environmental and Commercial Importance* (eds. K.N. Timmis and A. Puhler); Mercerau-Puigalon *et al.* (1980) *Gene* 11:163; Panthier *et al.* (1980) *Curr. Genet.* 2:109;).

A DNA molecule may be expressed intracellularly in yeast. A promoter sequence may be directly linked with the DNA molecule, in which case the first amino acid at the N-terminus of the recombinant protein will always be a methionine, which is encoded by the ATG start codon. If desired, methionine at the N-terminus may be cleaved from the protein by *in vitro* incubation with cyanogen bromide.

Fusion proteins provide an alternative for yeast expression systems, as well as in mammalian, plant, baculovirus, and bacterial expression systems. Usually, a DNA sequence encoding the N-terminal portion of an endogenous yeast protein, or other stable protein, is fused to the 5' end of heterologous coding sequences. Upon expression, this construct will provide a fusion of the two amino acid sequences. For example, the yeast or human

superoxide dismutase (SOD) gene, can be linked at the 5' terminus of a foreign gene and expressed in yeast. The DNA sequence at the junction of the two amino acid sequences may or may not encode a cleavable site. See e.g., EPO Publ. No. 196056. Another example is a ubiquitin fusion protein. Such a fusion protein is made with the ubiquitin region that preferably retains a site for a processing enzyme (e.g. ubiquitin-specific processing protease) to cleave the ubiquitin from the foreign protein. Through this method, therefore, native foreign protein can be isolated (e.g., WO88/024066).

Alternatively, foreign proteins can also be secreted from the cell into the growth media by creating chimeric DNA molecules that encode a fusion protein comprised of a leader sequence fragment that provide for secretion in yeast of the foreign protein. Preferably, there are processing sites encoded between the leader fragment and the foreign gene that can be cleaved either *in vivo* or *in vitro*. The leader sequence fragment usually encodes a signal peptide comprised of hydrophobic amino acids which direct the secretion of the protein from the cell.

DNA encoding suitable signal sequences can be derived from genes for secreted yeast proteins, such as the yeast invertase gene (EPO Publ. No. 012 873; JPO Publ. No. 62:096,086) and the A-factor gene (U.S. Patent 4,588,684). Alternatively, leaders of non-yeast origin, such as an interferon leader, exist that also provide for secretion in yeast (EPO Publ. No. 060 057).

A preferred class of secretion leaders are those that employ a fragment of the yeast alpha-factor gene, which contains both a "pre" signal sequence, and a "pro" region. The types of alpha-factor fragments that can be employed include the full-length pre-pro alpha factor leader (about 83 amino acid residues) as well as truncated alpha-factor leaders (usually about 25 to about 50 amino acid residues) (U.S. Patent Nos. 4,546,083 and 4,870,008; EPO Publ. No. 324 274). Additional leaders employing an alpha-factor leader fragment that provides for secretion include hybrid alpha-factor leaders made with a presequence of a first yeast, but a pro-region from a second yeast alphafactor. (See e.g., PCT Publ. No. WO 89/02463.)

Usually, transcription termination sequences recognized by yeast are regulatory regions located 3' to the translation stop codon, and thus together with the promoter flank the coding sequence. These sequences direct the transcription of an mRNA which can be translated into the polypeptide encoded by the DNA. Examples of transcription terminator

sequence and other yeast-recognized termination sequences, such as those coding for glycolytic enzymes.

Usually, the above described components, comprising a promoter, leader (if desired), coding sequence of interest, and transcription termination sequence, are put together into expression constructs. Expression constructs are often maintained in a replicon, such as an extrachromosomal element (e.g., plasmids) capable of stable maintenance in a host, such as yeast or bacteria. The replicon may have two replication systems, thus allowing it to be maintained, for example, in yeast for expression and in a prokaryotic host for cloning and amplification. Examples of such yeast-bacteria shuttle vectors include YEp24 (Botstein *et al.* (1979) *Gene* 8:17-24), pCl/1 (Brake *et al.* (1984) *Proc. Natl. Acad. Sci USA* 81:4642-4646), and YRp17 (Stinchcomb *et al.* (1982) *J. Mol. Biol.* 158:157). In addition, a replicon may be either a high or low copy number plasmid. A high copy number plasmid will generally have a copy number ranging from about 5 to about 200, and usually about 10 to about 150. A host containing a high copy number plasmid will preferably have at least about 10, and more preferably at least about 20. Enter a high or low copy number vector may be selected, depending upon the effect of the vector and the foreign protein on the host. See e.g., Brake *et al.*, *supra*.

Alternatively, the expression constructs can be integrated into the yeast genome with an integrating vector. Integrating vectors usually contain at least one sequence homologous to a yeast chromosome that allows the vector to integrate, and preferably contain two homologous sequences flanking the expression construct. Integrations appear to result from recombinations between homologous DNA in the vector and the yeast chromosome (Orr-Weaver *et al.* (1983) *Methods in Enzymol.* 101:228-245). An integrating vector may be directed to a specific locus in yeast by selecting the appropriate homologous sequence for inclusion in the vector. See Orr-Weaver *et al.*, *supra*. One or more expression construct may integrate, possibly affecting levels of recombinant protein produced (Rine *et al.* (1983) *Proc. Natl. Acad. Sci. USA* 80:6750). The chromosomal sequences included in the vector can occur either as a single segment in the vector, which results in the integration of the entire vector, or two segments homologous to adjacent segments in the chromosome and flanking the expression construct in the vector, which can result in the stable integration of only the expression construct.

Usually, extrachromosomal and integrating expression constructs may contain selectable markers to allow for the selection of yeast strains that have been transformed. Selectable markers may include biosynthetic genes that can be expressed in the yeast host, such as *ADE2*, *HIS4*, *LEU2*, *TRP1*, and *ALG7*, and the G418 resistance gene, which confer resistance in yeast cells to tunicamycin and G418, respectively. In addition, a suitable selectable marker may also provide yeast with the ability to grow in the presence of toxic compounds, such as metal. For example, the presence of *CUP1* allows yeast to grow in the presence of copper ions (Butt *et al.* (1987) *Microbiol. Rev.* 51:351).

Alternatively, some of the above described components can be put together into transformation vectors. Transformation vectors are usually comprised of a selectable marker that is either maintained in a replicon or developed into an integrating vector, as described above.

Expression and transformation vectors, either extrachromosomal replicons or integrating vectors, have been developed for transformation into many yeasts. For example, expression vectors and methods of introducing exogenous DNA into yeast hosts have been developed for, *inter alia*, the following yeasts: *Candida albicans* (Kurtz, *et al.* (1986) *Mol. Cell. Biol.* 6:142); *Candida maltosa* (Kunze, *et al.* (1985) *J. Basic Microbiol.* 25:141); *Hansenula polymorpha* (Gleeson, *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302); *Kluyveromyces fragilis* (Das, *et al.* (1984) *J. Bacteriol.* 158:1165); *Kluyveromyces lactis* (De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:737; Van den Berg *et al.* (1990) *Bio/Technology* 8:135); *Pichia guilliermondii* (Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141); *Pichia pastoris* (Cregg, *et al.* (1985) *Mol. Cell. Biol.* 5:3376; U.S. Patent Nos. 4,837,148 and 4,929,555); *Saccharomyces cerevisiae* (Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163); *Schizosaccharomyces pombe* (Beach and Nurse (1981) *Nature* 300:706); and *Yarrowia lipolytica* (Davidow, *et al.* (1985) *Curr. Genet.* 10:380471 Gaillardin, *et al.* (1985) *Curr. Genet.* 10:49).

Methods of introducing exogenous DNA into yeast hosts are well-known in the art, and usually include either the transformation of spheroplasts or of intact yeast cells treated with alkali cations. Transformation procedures usually vary with the yeast species to be transformed. See e.g., [Kurtz *et al.* (1986) *Mol. Cell. Biol.* 6:142; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; *Candida*]; [Gleeson *et al.* (1986) *J. Gen. Microbiol.* 132:3459; Roggenkamp *et al.* (1986) *Mol. Gen. Genet.* 202:302; *Hansenula*]; [Das *et al.* (1984) *J.*

Bacteriol. 158:1165; De Louvencourt *et al.* (1983) *J. Bacteriol.* 154:1165; Van den Berg *et al.* (1990) *Bio/Technology* 8:135; Kluyveromyces]; [Cregg *et al.* (1985) *Mol. Cell. Biol.* 5:3376; Kunze *et al.* (1985) *J. Basic Microbiol.* 25:141; U.S. Patent Nos. 4,837,148 and 4,929,555; Pichia]; [Hinnen *et al.* (1978) *Proc. Natl. Acad. Sci. USA* 75:1929; Ito *et al.* (1983) *J. Bacteriol.* 153:163 Saccharomyces]; [Beach and Nurse (1981) *Nature* 300:706; Schizosaccharomyces]; [Davidow *et al.* (1985) *Curr. Genet.* 10:39; Gaillardin *et al.* (1985) *Curr. Genet.* 10:49; Yarrowia].

Definitions

A composition containing X is "substantially free of" Y when at least 85% by weight of the total X+Y in the composition is X. Preferably, X comprises at least about 90% by weight of the total of X+Y in the composition, more preferably at least about 95% or even 99% by weight.

A "conserved" *Neisseria* amino acid fragment or protein is one that is present in a particular Neisserial protein in at least x% of *Neisseria*. The value of x may be 50% or more, e.g., 66%, 75%, 80%, 90%, 95% or even 100% (i.e. the amino acid is found in the protein in question in all *Neisseria*). In order to determine whether an amino acid is "conserved" in a particular Neisserial protein, it is necessary to compare that amino acid residue in the sequences of the protein in question from a plurality of different *Neisseria* (a reference population). The reference population may include a number of different *Neisseria* species or may include a single species. The reference population may include a number of different serogroups of a particular species or a single serogroup. A preferred reference population consists of the 5 most common *Neisseria* strains.

The term "heterologous" refers to two biological components that are not found together in nature. The components may be host cells, genes, or regulatory regions, such as promoters. Although the heterologous components are not found together in nature, they can function together, as when a promoter heterologous to a gene is operably linked to the gene. Another example is where a Neisserial sequence is heterologous to a mouse host cell.

"Epitope" means antigenic determinant, and may elicit a cellular and/or humoral response.

Conditions for "high stringency" are 65 degrees C in 0.1 xSSC 0.5% SDS solution.

An "origin of replication" is a polynucleotide sequence that initiates and regulates replication of polynucleotides, such as an expression vector. The origin of replication behaves as an autonomous unit of polynucleotide replication within a cell, capable of replication under its own control. An origin of replication may be needed for a vector to replicate in a particular host cell. With certain origins of replication, an expression vector can be reproduced at a high copy number in the presence of the appropriate proteins within the cell. Examples of origins are the autonomously replicating sequences, which are effective in yeast; and the viral T-antigen, effective in COS-7 cells.

A "mutant" sequence is defined as a DNA, RNA or amino acid sequence differing from but having homology with the native or disclosed sequence. Depending on the particular sequence, the degree of homology (sequence identity) between the native or disclosed sequence and the mutant sequence is preferably greater than 50% (e.g., 60%, 70%, 80%, 90%, 95%, 99% or more) which is calculated as described above. As used herein, an "allelic variant" of a nucleic acid molecule, or region, for which nucleic acid sequence is provided herein is a nucleic acid molecule, or region, that occurs at essentially the same locus in the genome of another or second isolate, and that, due to natural variation caused by, for example, mutation or recombination, has a similar but not identical nucleic acid sequence. A coding region allelic variant typically encodes a protein having similar activity to that of the protein encoded by the gene to which it is being compared. An allelic variant can also comprise an alteration in the 5' or 3' untranslated regions of the gene, such as in regulatory control regions. (see, for example, U.S. Patent 5,753,235).

Antibodies

As used herein, the term "antibody" refers to a polypeptide or group of polypeptides composed of at least one antibody combining site. An "antibody combining site" is the three-dimensional binding space with an internal surface shape and charge distribution complementary to the features of an epitope of an antigen, which allows a binding of the antibody with the antigen. "Antibody" includes, for example, vertebrate antibodies, hybrid antibodies, chimeric antibodies, humanized antibodies, altered antibodies, univalent antibodies, Fab proteins, and single domain antibodies.

Antibodies against the proteins of the invention are useful for affinity chromatography, immunoassays, and distinguishing/identifying *Neisseria* menB proteins.

Antibodies elicited against the proteins of the present invention bind to antigenic polypeptides or proteins or protein fragments that are present and specifically associated with strains of *Neisseria meningitidis* menB. In some instances, these antigens may be associated with specific strains, such as those antigens specific for the menB strains. The antibodies of the invention may be immobilized to a matrix and utilized in an immunoassay or on an affinity chromatography column, to enable the detection and/or separation of polypeptides, proteins or protein fragments or cells comprising such polypeptides, proteins or protein fragments. Alternatively, such polypeptides, proteins or protein fragments may be immobilized so as to detect antibodies bindably specific thereto.

Antibodies to the proteins of the invention, both polyclonal and monoclonal, may be prepared by conventional methods. In general, the protein is first used to immunize a suitable animal, preferably a mouse, rat, rabbit or goat. Rabbits and goats are preferred for the preparation of polyclonal sera due to the volume of serum obtainable, and the availability of labeled anti-rabbit and anti-goat antibodies. Immunization is generally performed by mixing or emulsifying the protein in saline, preferably in an adjuvant such as Freund's complete adjuvant, and injecting the mixture or emulsion parenterally (generally subcutaneously or intramuscularly). A dose of 50-200 µg/injection is typically sufficient. Immunization is generally boosted 2-6 weeks later with one or more injections of the protein in saline, preferably using Freund's incomplete adjuvant. One may alternatively generate antibodies by *in vitro* immunization using methods known in the art, which for the purposes of this invention is considered equivalent to *in vivo* immunization. Polyclonal antisera is obtained by bleeding the immunized animal into a glass or plastic container, incubating the blood at 25°C for one hour, followed by incubating at 4°C for 2-18 hours. The serum is recovered by centrifugation (e.g., 1,000g for 10 minutes). About 20-50 ml per bleed may be obtained from rabbits.

Monoclonal antibodies are prepared using the standard method of Kohler & Milstein (*Nature* (1975) 256:495-96), or a modification thereof. Typically, a mouse or rat is immunized as described above. However, rather than bleeding the animal to extract serum, the spleen (and optionally several large lymph nodes) is removed and dissociated into single cells. If desired, the spleen cells may be screened (after removal of nonspecifically adherent cells) by applying a cell suspension to a plate or well coated with the protein antigen. B-cells that express membrane-bound immunoglobulin specific for the antigen bind to the plate, and

are not rinsed away with the rest of the suspension. Resulting B-cells, or all dissociated spleen cells, are then induced to fuse with myeloma cells to form hybridomas, and are cultured in a selective medium (e.g., hypoxanthine, aminopterin, thymidine medium, "HAT"). The resulting hybridomas are plated by limiting dilution, and are assayed for the production of antibodies which bind specifically to the immunizing antigen (and which do not bind to unrelated antigens). The selected MAb-secreting hybridomas are then cultured either *in vitro* (e.g., in tissue culture bottles or hollow fiber reactors), or *in vivo* (as ascites in mice).

If desired, the antibodies (whether polyclonal or monoclonal) may be labeled using conventional techniques. Suitable labels include fluorophores, chromophores, radioactive atoms (particularly ^{32}P and ^{125}I), electron-dense reagents, enzymes, and ligands having specific binding partners. Enzymes are typically detected by their activity. For example, horseradish peroxidase is usually detected by its ability to convert 3,3',5,5'-tetramethylbenzidine (TMB) to a blue pigment, quantifiable with a spectrophotometer. "Specific binding partner" refers to a protein capable of binding a ligand molecule with high specificity, as for example in the case of an antigen and a monoclonal antibody specific therefor. Other specific binding partners include biotin and avidin or streptavidin, IgG and protein A, and the numerous receptor-ligand couples known in the art. It should be understood that the above description is not meant to categorize the various labels into distinct classes, as the same label may serve in several different modes. For example, ^{125}I may serve as a radioactive label or as an electron-dense reagent. HRP may serve as enzyme or as antigen for a MAb. Further, one may combine various labels for desired effect. For example, MAbs and avidin also require labels in the practice of this invention: thus, one might label a MAb with biotin, and detect its presence with avidin labeled with ^{125}I , or with an anti-biotin MAb labeled with HRP. Other permutations and possibilities will be readily apparent to those of ordinary skill in the art, and are considered as equivalents within the scope of the instant invention.

Antigens, immunogens, polypeptides, proteins or protein fragments of the present invention elicit formation of specific binding partner antibodies. These antigens, immunogens, polypeptides, proteins or protein fragments of the present invention comprise immunogenic compositions of the present invention. Such immunogenic compositions may further comprise or include adjuvants, carriers, or other compositions that promote or enhance

or stabilize the antigens, polypeptides, proteins or protein fragments of the present invention. Such adjuvants and carriers will be readily apparent to those of ordinary skill in the art.

Pharmaceutical Compositions

Pharmaceutical compositions can comprise (include) either polypeptides, antibodies, or nucleic acid of the invention. The pharmaceutical compositions will comprise a therapeutically effective amount of either polypeptides, antibodies, or polynucleotides of the claimed invention.

The term "therapeutically effective amount" as used herein refers to an amount of a therapeutic agent to treat, ameliorate, or prevent a desired disease or condition, or to exhibit a detectable therapeutic or preventative effect. The effect can be detected by, for example, chemical markers or antigen levels. Therapeutic effects also include reduction in physical symptoms, such as decreased body temperature, when given to a patient that is febrile. The precise effective amount for a subject will depend upon the subject's size and health, the nature and extent of the condition, and the therapeutics or combination of therapeutics selected for administration. Thus, it is not useful to specify an exact effective amount in advance. However, the effective amount for a given situation can be determined by routine experimentation and is within the judgment of the clinician.

For purposes of the present invention, an effective dose will be from about 0.01 mg/kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

A pharmaceutical composition can also contain a pharmaceutically acceptable carrier. The term "pharmaceutically acceptable carrier" refers to a carrier for administration of a therapeutic agent, such as antibodies or a polypeptide, genes, and other therapeutic agents. The term refers to any pharmaceutical carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition, and which may be administered without undue toxicity. Suitable carriers may be large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, and inactive virus particles. Such carriers are well known to those of ordinary skill in the art.

Pharmaceutically acceptable salts can be used therein, for example, mineral acid salts such as hydrochlorides, hydrobromides, phosphates, sulfates, and the like; and the salts of

organic acids such as acetates, propionates, malonates, benzoates, and the like. A thorough discussion of pharmaceutically acceptable excipients is available in Remington's Pharmaceutical Sciences (Mack Pub. Co., N.J. 1991).

Pharmaceutically acceptable carriers in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Typically, the therapeutic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. Liposomes are included within the definition of a pharmaceutically acceptable carrier.

Delivery Methods

Once formulated, the compositions of the invention can be administered directly to the subject. The subjects to be treated can be animals; in particular, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal and transcutaneous applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Vaccines

Vaccines according to the invention may either be prophylactic (i.e., to prevent infection) or therapeutic (i.e., to treat disease after infection).

Such vaccines comprise immunizing antigen(s) or immunogen(s), immunogenic polypeptide, protein(s) or protein fragments, or nucleic acids (e.g., ribonucleic acid or deoxyribonucleic acid), usually in combination with "pharmaceutically acceptable carriers," which include any carrier that does not itself induce the production of antibodies harmful to the individual receiving the composition. Suitable carriers are typically large, slowly metabolized macromolecules such as proteins, polysaccharides, polylactic acids, polyglycolic acids, polymeric amino acids, amino acid copolymers, lipid aggregates (such as oil droplets or

liposomes), and inactive virus particles. Such carriers are well known to those of ordinary skill in the art. Additionally, these carriers may function as immunostimulating agents ("adjuvants"). Furthermore, the immunogen or antigen may be conjugated to a bacterial toxoid, such as a toxoid from diphtheria, tetanus, cholera, *H. pylori*, etc. pathogens.

Preferred adjuvants to enhance effectiveness of the composition include, but are not limited to: (1) aluminum salts (alum), such as aluminum hydroxide, aluminum phosphate, aluminum sulfate, etc; (2) oil-in-water emulsion formulations (with or without other specific immunostimulating agents such as muramyl peptides (see below) or bacterial cell wall components), such as for example (a) MF59 (PCT Publ. No. WO 90/14837), containing 5% Squalene, 0.5% Tween 80, and 0.5% Span 85 (optionally containing various amounts of MTP-PE (see below), although not required) formulated into submicron particles using a microfluidizer such as Model 110Y microfluidizer (Microfluidics, Newton, MA), (b) SAF, containing 10% Squalene, 0.4% Tween 80, 5% pluronic-blocked polymer L121, and thr-MDP (see below) either microfluidized into a submicron emulsion or vortexed to generate a larger particle size emulsion, and (c) RibiTM adjuvant system (RAS), (Ribi Immunochem, Hamilton, MT) containing 2% Squalene, 0.2% Tween 80, and one or more bacterial cell wall components from the group consisting of monophosphorylipid A (MPL), trehalose dimycolate (TDM), and cell wall skeleton (CWS), preferably MPL + CWS (DetoxTM); (3) saponin adjuvants, such as StimulonTM (Cambridge Bioscience, Worcester, MA) may be used or particles generated therefrom such as ISCOMs (immunostimulating complexes); (4) Complete Freund's Adjuvant (CFA) and Incomplete Freund's Adjuvant (IFA); (5) cytokines, such as interleukins (e.g., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferons (e.g., gamma interferon), macrophage colony stimulating factor (M-CSF), tumor necrosis factor (TNF), etc; (6) detoxified mutants of a bacterial ADP-ribosylating toxin such as a cholera toxin (CT), a pertussis toxin (PT), or an *E. coli* heat-labile toxin (LT), particularly LT-K63, LT-R72, CT-S109, PT-K9/G129; see, e.g., WO 93/13302 and WO 92/19265; and (7) other substances that act as immunostimulating agents to enhance the effectiveness of the composition. Alum and MF59 are preferred.

As mentioned above, muramyl peptides include, but are not limited to, N-acetyl-muramyl-L-threonyl-D-isoglutamine (thr-MDP), N-acetyl-normuramyl-L-alanyl-D-isoglutamine (nor-MDP), N-acetylmuramyl-L-alanyl-D-isoglutaminyl-L-alanine-2-(1'-2'-dipalmitoyl-*sn*-glycero-3-hydroxyphosphoryloxy)-ethylamine (MTP-PE), etc.

The vaccine compositions comprising immunogenic compositions (e.g., which may include the antigen, pharmaceutically acceptable carrier, and adjuvant) typically will contain diluents, such as water, saline, glycerol, ethanol, etc. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles. Alternatively, vaccine compositions comprising immunogenic compositions may comprise an antigen, polypeptide, protein, protein fragment or nucleic acid in a pharmaceutically acceptable carrier.

More specifically, vaccines comprising immunogenic compositions comprise an immunologically effective amount of the immunogenic polypeptides, as well as any other of the above-mentioned components, as needed. By "immunologically effective amount", it is meant that the administration of that amount to an individual, either in a single dose or as part of a series, is effective for treatment or prevention. This amount varies depending upon the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated (e.g., nonhuman primate, primate, *etc.*), the capacity of the individual's immune system to synthesize antibodies, the degree of protection desired, the formulation of the vaccine, the treating doctor's assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

Typically, the vaccine compositions or immunogenic compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid vehicles prior to injection may also be prepared. The preparation also may be emulsified or encapsulated in liposomes for enhanced adjuvant effect, as discussed above under pharmaceutically acceptable carriers.

The immunogenic compositions are conventionally administered parenterally, e.g., by injection, either subcutaneously or intramuscularly. Additional formulations suitable for other modes of administration include oral and pulmonary formulations, suppositories, and transdermal and transcutaneous applications. Dosage treatment may be a single dose schedule or a multiple dose schedule. The vaccine may be administered in conjunction with other immunoregulatory agents.

As an alternative to protein-based vaccines, DNA vaccination may be employed (e.g., Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly *et al.* (1997) *Annu Rev Immunol* 15:617-648).

Gene Delivery Vehicles

Gene therapy vehicles for delivery of constructs, including a coding sequence of a therapeutic of the invention, to be delivered to the mammal for expression in the mammal, can be administered either locally or systemically. These constructs can utilize viral or non-viral vector approaches in *in vivo* or *ex vivo* modality. Expression of such coding sequence can be induced using endogenous mammalian or heterologous promoters. Expression of the coding sequence in vivo can be either constitutive or regulated.

The invention includes gene delivery vehicles capable of expressing the contemplated nucleic acid sequences. The gene delivery vehicle is preferably a viral vector and, more preferably, a retroviral, adenoviral, adeno-associated viral (AAV), herpes viral, or alphavirus vector. The viral vector can also be an astrovirus, coronavirus, orthomyxovirus, papovavirus, paramyxovirus, parvovirus, picornavirus, poxvirus, or togavirus viral vector. See generally, Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; and Kaplitt (1994) *Nature Genetics* 6:148-153.

Retroviral vectors are well known in the art, including B, C and D type retroviruses, xenotropic retroviruses (for example, NZB-X1, NZB-X2 and NZB9-1 (see O'Neill (1985) *J. Virol.* 53:160) polytropic retroviruses e.g., MCF and MCF-MLV (see Kelly (1983) *J. Virol.* 45:291), spumaviruses and lentiviruses. See RNA Tumor Viruses, Second Edition, Cold Spring Harbor Laboratory, 1985.

Portions of the retroviral gene therapy vector may be derived from different retroviruses. For example, retrovector LTRs may be derived from a Murine Sarcoma Virus, a tRNA binding site from a Rous Sarcoma Virus, a packaging signal from a Murine Leukemia Virus, and an origin of second strand synthesis from an Avian Leukosis Virus.

These recombinant retroviral vectors may be used to generate transduction competent retroviral vector particles by introducing them into appropriate packaging cell lines (see US patent 5,591,624). Retrovirus vectors can be constructed for site-specific integration into host cell DNA by incorporation of a chimeric integrase enzyme into the retroviral particle (see WO96/37626). It is preferable that the recombinant viral vector is a replication defective recombinant virus.

Packaging cell lines suitable for use with the above-described retrovirus vectors are well known in the art, are readily prepared (see WO95/30763 and WO92/05266), and can be used to create producer cell lines (also termed vector cell lines or "VCLs") for the production of recombinant vector particles. Preferably, the packaging cell lines are made from human parent cells (e.g., HT1080 cells) or mink parent cell lines, which eliminates inactivation in human serum.

Preferred retroviruses for the construction of retroviral gene therapy vectors include Avian Leukosis Virus, Bovine Leukemia Virus, Murine Leukemia Virus, Mink-Cell Focus-Inducing Virus, Murine Sarcoma Virus, Reticuloendotheliosis Virus and Rous Sarcoma Virus. Particularly preferred Murine Leukemia Viruses include 4070A and 1504A (Hartley and Rowe (1976) *J Virol* 19:19-25), Abelson (ATCC No. VR-999), Friend (ATCC No. VR-245), Graffi, Gross (ATCC No. VR-590), Kirsten, Harvey Sarcoma Virus and Rauscher (ATCC No. VR-998) and Moloney Murine Leukemia Virus (ATCC No. VR-190). Such retroviruses may be obtained from depositories or collections such as the American Type Culture Collection ("ATCC") in Rockville, Maryland or isolated from known sources using commonly available techniques.

Exemplary known retroviral gene therapy vectors employable in this invention include those described in patent applications GB2200651, EP0415731, EP0345242, EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622, WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. See also Vile (1993) *Cancer Res* 53:3860-3864; Vile (1993) *Cancer Res* 53:962-967; Ram (1993) *Cancer Res* 53 (1993) 83-88; Takamiya (1992) *J Neurosci Res* 33:493-503; Baba (1993) *J Neurosurg* 79:729-735; Mann (1983) *Cell* 33:153; Cane (1984) *Proc Natl Acad Sci* 81:6349; and Miller (1990) *Human Gene Therapy* 1.

Human adenoviral gene therapy vectors are also known in the art and employable in this invention. See, for example, Berkner (1988) *Biotechniques* 6:616 and Rosenfeld (1991) *Science* 252:431, and WO93/07283, WO93/06223, and WO93/07282. Exemplary known adenoviral gene therapy vectors employable in this invention include those described in the above referenced documents and in WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102,

WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 and WO95/09654. Alternatively, administration of DNA linked to killed adenovirus as described in Curiel (1992) *Hum. Gene Ther.* 3:147-154 may be employed. The gene delivery vehicles of the invention also include adenovirus associated virus (AAV) vectors. Leading and preferred examples of such vectors for use in this invention are the AAV-2 based vectors disclosed in Srivastava, WO93/09239. Most preferred AAV vectors comprise the two AAV inverted terminal repeats in which the native D-sequences are modified by substitution of nucleotides, such that at least 5 native nucleotides and up to 18 native nucleotides, preferably at least 10 native nucleotides up to 18 native nucleotides, most preferably 10 native nucleotides are retained and the remaining nucleotides of the D-sequence are deleted or replaced with non-native nucleotides. The native D-sequences of the AAV inverted terminal repeats are sequences of 20 consecutive nucleotides in each AAV inverted terminal repeat (i.e., there is one sequence at each end) which are not involved in HP formation. The non-native replacement nucleotide may be any nucleotide other than the nucleotide found in the native D-sequence in the same position. Other employable exemplary AAV vectors are pWP-19, pWN-1, both of which are disclosed in Nahreini (1993) *Gene* 124:257-262. Another example of such an AAV vector is psub201 (see Samulski (1987) *J. Virol.* 61:3096). Another exemplary AAV vector is the Double-D ITR vector. Construction of the Double-D ITR vector is disclosed in US Patent 5,478,745. Still other vectors are those disclosed in Carter US Patent 4,797,368 and Muzyczka US Patent 5,139,941, Chartejee US Patent 5,474,935, and Kotin WO94/288157. Yet a further example of an AAV vector employable in this invention is SSV9AFABTKneo, which contains the AFP enhancer and albumin promoter and directs expression predominantly in the liver. Its structure and construction are disclosed in Su (1996) *Human Gene Therapy* 7:463-470. Additional AAV gene therapy vectors are described in US 5,354,678, US 5,173,414, US 5,139,941, and US 5,252,479.

The gene therapy vectors comprising sequences of the invention also include herpes vectors. Leading and preferred examples are herpes simplex virus vectors containing a sequence encoding a thymidine kinase polypeptide such as those disclosed in US 5,288,641 and EP0176170 (Roizman). Additional exemplary herpes simplex virus vectors include HFEM/ICP6-LacZ disclosed in WO95/04139 (Wistar Institute), pHSVlac described in Geller (1988) *Science* 241:1667-1669 and in WO90/09441 and WO92/07945, HSV Us3::pgC-lacZ

described in Fink (1992) *Human Gene Therapy* 3:11-19 and HSV 7134, 2 RH 105 and GAL4 described in EP 0453242 (Breakefield), and those deposited with the ATCC as accession numbers ATCC VR-977 and ATCC VR-260.

Also contemplated are alpha virus gene therapy vectors that can be employed in this invention. Preferred alpha virus vectors are Sindbis viruses vectors. Togaviruses, Semliki Forest virus (ATCC VR-67; ATCC VR-1247), Middleberg virus (ATCC VR-370), Ross River virus (ATCC VR-373; ATCC VR-1246), Venezuelan equine encephalitis virus (ATCC VR923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532), and those described in US patents 5,091,309, 5,217,879, and WO92/10578. More particularly, those alpha virus vectors described in U.S. Serial No. 08/405,627, filed March 15, 1995, WO94/21792, WO92/10578, WO95/07994, US 5,091,309 and US 5,217,879 are employable. Such alpha viruses may be obtained from depositories or collections such as the ATCC in Rockville, Maryland or isolated from known sources using commonly available techniques. Preferably, alphavirus vectors with reduced cytotoxicity are used (see USSN 08/679640).

DNA vector systems such as eukaryotic layered expression systems are also useful for expressing the nucleic acids of the invention. See WO95/07994 for a detailed description of eukaryotic layered expression systems. Preferably, the eukaryotic layered expression systems of the invention are derived from alphavirus vectors and most preferably from Sindbis viral vectors.

Other viral vectors suitable for use in the present invention include those derived from poliovirus, for example ATCC VR-58 and those described in Evans, *Nature* 339 (1989) 385 and Sabin (1973) *J. Biol. Standardization* 1:115; rhinovirus, for example ATCC VR-1110 and those described in Arnold (1990) *J Cell Biochem* L401; pox viruses such as canary pox virus or vaccinia virus, for example ATCC VR-111 and ATCC VR-2010 and those described in Fisher-Hoch (1989) *Proc Natl Acad Sci* 86:317; Flexner (1989) *Ann NY Acad Sci* 569:86, Flexner (1990) *Vaccine* 8:17; in US 4,603,112 and US 4,769,330 and WO89/01973; SV40 virus, for example ATCC VR-305 and those described in Mulligan (1979) *Nature* 277:108 and Madzak (1992) *J Gen Virol* 73:1533; influenza virus, for example ATCC VR-797 and recombinant influenza viruses made employing reverse genetics techniques as described in US 5,166,057 and in Enami (1990) *Proc Natl Acad Sci* 87:3802-3805; Enami & Palese (1991) *J Virol* 65:2711-2713 and Luytjes (1989) *Cell* 59:110, (see also McMichael (1983) *NEJ Med* 309:13, and Yap (1978) *Nature* 273:238 and *Nature* (1979) 277:108); human

immunodeficiency virus as described in EP-0386882 and in Buchschacher (1992) *J. Virol.* 66:2731; measles virus, for example ATCC VR-67 and VR-1247 and those described in EP-0440219; Aura virus, for example ATCC VR-368; Bebaru virus, for example ATCC VR-600 and ATCC VR-1240; Cabassou virus, for example ATCC VR-922; Chikungunya virus, for example ATCC VR-64 and ATCC VR-1241; Fort Morgan Virus, for example ATCC VR-924; Getah virus, for example ATCC VR-369 and ATCC VR-1243; Kyzylagach virus, for example ATCC VR-927; Mayaro virus, for example ATCC VR-66; Mucambo virus, for example ATCC VR-580 and ATCC VR-1244; Ndumu virus, for example ATCC VR-371; Pixuna virus, for example ATCC VR-372 and ATCC VR-1245; Tonate virus, for example ATCC VR-925; Trinita virus, for example ATCC VR-469; Una virus, for example ATCC VR-374; Whataroa virus, for example ATCC VR-926; Y-62-33 virus, for example ATCC VR-375; O'Nyong virus, Eastern encephalitis virus, for example ATCC VR-65 and ATCC VR-1242; Western encephalitis virus, for example ATCC VR-70, ATCC VR-1251, ATCC VR-622 and ATCC VR-1252; and coronavirus, for example ATCC VR-740 and those described in Hamre (1966) *Proc Soc Exp Biol Med* 121:190.

Delivery of the compositions of this invention into cells is not limited to the above mentioned viral vectors. Other delivery methods and media may be employed such as, for example, nucleic acid expression vectors, polycationic condensed DNA linked or unlinked to killed adenovirus alone, for example see US Serial No. 08/366,787, filed December 30, 1994 and Curiel (1992) *Hum Gene Ther* 3:147-154 ligand linked DNA, for example see Wu (1989) *J Biol Chem* 264:16985-16987, eucaryotic cell delivery vehicles cells, for example see US Serial No.08/240,030, filed May 9, 1994, and US Serial No. 08/404,796, deposition of photopolymerized hydrogel materials, hand-held gene transfer particle gun, as described in US Patent 5,149,655, ionizing radiation as described in US5,206,152 and in WO92/11033, nucleic charge neutralization or fusion with cell membranes. Additional approaches are described in Philip (1994) *Mol Cell Biol* 14:2411-2418 and in Woffendin (1994) *Proc Natl Acad Sci* 91:1581-1585.

Particle mediated gene transfer may be employed, for example see US Serial No. 60/023,867. Briefly, the sequence can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, as described in Wu &

Wu (1987) *J. Biol. Chem.* 262:4429-4432, insulin as described in Hucked (1990) *Biochem Pharmacol* 40:253-263, galactose as described in Plank (1992) *Bioconjugate Chem* 3:533-539, lactose or transferrin.

Naked DNA may also be employed to transform a host cell. Exemplary naked DNA introduction methods are described in WO 90/11092 and US 5,580,859. Uptake efficiency -- may be improved using biodegradable latex beads. DNA coated latex beads are efficiently transported into cells after endocytosis initiation by the beads. The method may be improved further by treatment of the beads to increase hydrophobicity and thereby facilitate disruption of the endosome and release of the DNA into the cytoplasm.

Liposomes that can act as gene delivery vehicles are described in U.S. 5,422,120, WO95/13796, WO94/23697, WO91/14445 and EP-524,968. As described in USSN. 60/023,867, on non-viral delivery, the nucleic acid sequences encoding a polypeptide can be inserted into conventional vectors that contain conventional control sequences for high level expression, and then be incubated with synthetic gene transfer molecules such as polymeric DNA-binding cations like polylysine, protamine, and albumin, linked to cell targeting ligands such as asialoorosomucoid, insulin, galactose, lactose, or transferrin. Other delivery systems include the use of liposomes to encapsulate DNA comprising the gene under the control of a variety of tissue-specific or ubiquitously-active promoters. Further non-viral delivery suitable for use includes mechanical delivery systems such as the approach described in Woffendin *et al* (1994) *Proc. Natl. Acad. Sci. USA* 91(24):11581-11585. Moreover, the coding sequence and the product of expression of such can be delivered through deposition of photopolymerized hydrogel materials. Other conventional methods for gene delivery that can be used for delivery of the coding sequence include, for example, use of hand-held gene transfer particle gun, as described in U.S. 5,149,655; use of ionizing radiation for activating transferred gene, as described in U.S. 5,206,152 and WO92/11033.

Exemplary liposome and polycationic gene delivery vehicles are those described in US 5,422,120 and 4,762,915; in WO 95/13796; WO94/23697; and WO91/14445; in EP- 0524968; and in Stryer, *Biochemistry*, pages 236-240 (1975) W.H. Freeman, San Francisco; Szoka (1980) *Biochem Biophys Acta* 600:1; Bayer (1979) *Biochem Biophys Acta* 550:464; Rivnay (1987) *Meth Enzymol* 149:119; Wang (1987) *Proc Natl Acad Sci* 84:7851; Plant (1989) *Anal Biochem* 176:420.

A polynucleotide composition can comprises therapeutically effective amount of a gene therapy vehicle, as the term is defined above. For purposes of the present invention, an effective dose will be from about 0.01 mg/ kg to 50 mg/kg or 0.05 mg/kg to about 10 mg/kg of the DNA constructs in the individual to which it is administered.

Delivery Methods

Once formulated, the polynucleotide compositions of the invention can be administered (1) directly to the subject; (2) delivered *ex vivo*, to cells derived from the subject; or (3) *in vitro* for expression of recombinant proteins. The subjects to be treated can be mammals or birds. Also, human subjects can be treated.

Direct delivery of the compositions will generally be accomplished by injection, either subcutaneously, intraperitoneally, intravenously or intramuscularly or delivered to the interstitial space of a tissue. The compositions can also be administered into a tumor or lesion. Other modes of administration include oral and pulmonary administration, suppositories, and transdermal applications, needles, and gene guns or hyposprays. Dosage treatment may be a single dose schedule or a multiple dose schedule.

Methods for the *ex vivo* delivery and reimplantation of transformed cells into a subject are known in the art and described in *eg.* WO93/14778. Examples of cells useful in *ex vivo* applications include, for example, stem cells, particularly hematopoietic, lymph cells, macrophages, dendritic cells, or tumor cells.

Generally, delivery of nucleic acids for both *ex vivo* and *in vitro* applications can be accomplished by the following procedures, for example, dextran-mediated transfection, calcium phosphate precipitation, polybrene mediated transfection, protoplast fusion, electroporation, encapsulation of the polynucleotide(s) in liposomes, and direct microinjection of the DNA into nuclei, all well known in the art.

Polynucleotide and polypeptide pharmaceutical compositions

In addition to the pharmaceutically acceptable carriers and salts described above, the following additional agents can be used with polynucleotide and/or polypeptide compositions.

A. Polypeptides

One example are polypeptides which include, without limitation: asioloorosomucoid (ASOR); transferrin; asialoglycoproteins; antibodies; antibody fragments; ferritin; interleukins; interferons, granulocyte, macrophage colony stimulating factor (GM-CSF),

granulocyte colony stimulating factor (G-CSF), macrophage colony stimulating factor (M-CSF), stem cell factor and erythropoietin. Viral antigens, such as envelope proteins, can also be used. Also, proteins from other invasive organisms, such as the 17 amino acid peptide from the circumsporozoite protein of plasmodium falciparum known as RII.

B.Hormones, Vitamins, Etc.

Other groups that can be included are, for example: hormones, steroids, androgens, estrogens, thyroid hormone, or vitamins, folic acid.

C.Polyalkylenes, Polysaccharides, etc.

Also, polyalkylene glycol can be included with the desired polynucleotides or polypeptides. In a preferred embodiment, the polyalkylene glycol is polyethylene glycol. In addition, mono-, di-, or polysaccharides can be included. In a preferred embodiment of this aspect, the polysaccharide is dextran or DEAE-dextran. Also, chitosan and poly(lactide-co-glycolide)

D.Lipids, and Liposomes

The desired polynucleotide or polypeptide can also be encapsulated in lipids or packaged in liposomes prior to delivery to the subject or to cells derived therefrom.

Lipid encapsulation is generally accomplished using liposomes which are able to stably bind or entrap and retain nucleic acid. The ratio of condensed polynucleotide or polypeptide to lipid preparation can vary but will generally be around 1:1 (mg DNA:micromoles lipid), or more of lipid. For a review of the use of liposomes as carriers for delivery of nucleic acids, see, Hug and Sleight (1991) *Biochim. Biophys. Acta.* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

Liposomal preparations for use in the present invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); and purified transcription factors (Debs (1990) *J. Biol. Chem.* 265:10189-10192), in functional form.

Cationic liposomes are readily available. For example, N[1-2,3-dioleoyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, NY. (See, also, Felgner *supra*). Other commercially available liposomes include transfectace (DDAB/DOPE) and

DOTAP/DOPE (Boehringer). Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, eg. Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; WO90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes.

Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, AL), or can be easily prepared using readily available materials. Such materials include phosphatidyl choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphosphatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs). The various liposome-nucleic acid complexes are prepared using methods known in the art. See eg. Straubinger (1983) *Meth. Immunol.* 101:512-527; Szoka (1978) *Proc. Natl. Acad. Sci. USA* 75:4194-4198; Papahadjopoulos (1975) *Biochim. Biophys. Acta* 394:483; Wilson (1979) *Cell* 17:77; Deamer & Bangham (1976) *Biochim. Biophys. Acta* 443:629; Ostro (1977) *Biochem. Biophys. Res. Commun.* 76:836; Fraley (1979) *Proc. Natl. Acad. Sci. USA* 76:3348; Enoch & Strittmatter (1979) *Proc. Natl. Acad. Sci. USA* 76:145; Fraley (1980) *J. Biol. Chem.* (1980) 255:10431; Szoka & Papahadjopoulos (1978) *Proc. Natl. Acad. Sci. USA* 75:145; and Schaefer-Ridder (1982) *Science* 215:166.

E.Lipoproteins

In addition, lipoproteins can be included with the polynucleotide or polypeptide to be delivered. Examples of lipoproteins to be utilized include: chylomicrons, HDL, IDL, LDL, and VLDL. Mutants, fragments, or fusions of these proteins can also be used. Also, modifications of naturally occurring lipoproteins can be used, such as acetylated LDL. These lipoproteins can target the delivery of polynucleotides to cells expressing lipoprotein receptors. Preferably, if lipoproteins are including with the polynucleotide to be delivered, no other targeting ligand is included in the composition.

Naturally occurring lipoproteins comprise a lipid and a protein portion. The protein portion are known as apoproteins. At the present, apoproteins A, B, C, D, and E have been

isolated and identified. At least two of these contain several proteins, designated by Roman numerals, AI, AII, AIV; CI, CII, CIII.

A lipoprotein can comprise more than one apoprotein. For example, naturally occurring chylomicrons comprises of A, B, C, and E, over time these lipoproteins lose A and acquire C and E apoproteins. VLDL comprises A, B, C, and E apoproteins, LDL comprises apoprotein B; and HDL comprises apoproteins A, C, and E.

The amino acid of these apoproteins are known and are described in, for example, Breslow (1985) *Annu Rev. Biochem* 54:699; Law (1986) *Adv. Exp Med. Biol.* 151:162; Chen (1986) *J Biol Chem* 261:12918; Kane (1980) *Proc Natl Acad Sci USA* 77:2465; and Utermann (1984) *Hum Genet* 65:232.

Lipoproteins contain a variety of lipids including, triglycerides, cholesterol (free and esters), and phospholipids. The composition of the lipids varies in naturally occurring lipoproteins. For example, chylomicrons comprise mainly triglycerides. A more detailed description of the lipid content of naturally occurring lipoproteins can be found, for example, in *Meth. Enzymol.* 128 (1986). The composition of the lipids are chosen to aid in conformation of the apoprotein for receptor binding activity. The composition of lipids can also be chosen to facilitate hydrophobic interaction and association with the polynucleotide binding molecule.

Naturally occurring lipoproteins can be isolated from serum by ultracentrifugation, for instance. Such methods are described in *Meth. Enzymol.* (*supra*); Pitas (1980) *J. Biochem.* 255:5454-5460 and Mahey (1979) *J Clin. Invest* 64:743-750.

Lipoproteins can also be produced by *in vitro* or recombinant methods by expression of the apoprotein genes in a desired host cell. See, for example, Atkinson (1986) *Annu Rev Biophys Chem* 15:403 and Radding (1958) *Biochim Biophys Acta* 30: 443.

Lipoproteins can also be purchased from commercial suppliers, such as Biomedical Technologies, Inc., Stoughton, Massachusetts, USA.

Further description of lipoproteins can be found in Zuckermann et al., PCT. Appln. No. US97/14465.

F. Polycationic Agents

Polycationic agents can be included, with or without lipoprotein, in a composition with the desired polynucleotide or polypeptide to be delivered.

Polycationic agents, typically, exhibit a net positive charge at physiological relevant pH and are capable of neutralizing the electrical charge of nucleic acids to facilitate delivery to a desired location. These agents have both in vitro, ex vivo, and in vivo applications. Polycationic agents can be used to deliver nucleic acids to a living subject either intramuscularly, subcutaneously, etc.

The following are examples of useful polypeptides as polycationic agents: polylysine, polyarginine, polyornithine, and protamine. Other examples include histones, protamines, human serum albumin, DNA binding proteins, non-histone chromosomal proteins, coat proteins from DNA viruses, such as (X174, transcriptional factors also contain domains that bind DNA and therefore may be useful as nucleic acid condensing agents. Briefly, transcriptional factors such as C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP, and TFIID contain basic domains that bind DNA sequences.

Organic polycationic agents include: spermine, spermidine, and putrescine.

The dimensions and of the physical properties of a polycationic agent can be extrapolated from the list above, to construct other polypeptide polycationic agents or to produce synthetic polycationic agents.

Synthetic Polycationic Agents

Synthetic polycationic agents which are useful include, for example, DEAE-dextran, polybrene. Lipofectin[®], and lipofectAMINE[®] are monomers that form polycationic complexes when combined with polynucleotides or polypeptides.

Immunodiagnostic Assays

Neisserial antigens of the invention can be used in immunoassays to detect antibody levels (or, conversely, anti-Neisserial antibodies can be used to detect antigen levels). Immunoassays based on well defined, recombinant antigens can be developed to replace invasive diagnostics methods. Antibodies to Neisserial proteins within biological samples, including for example, blood or serum samples, can be detected. Design of the immunoassays is subject to a great deal of variation, and a variety of these are known in the art. Protocols for the immunoassay may be based, for example, upon competition, or direct reaction, or sandwich type assays. Protocols may also, for example, use solid supports, or may be by immunoprecipitation. Most assays involve the use of labeled antibody or polypeptide; the labels may be, for example, fluorescent, chemiluminescent, radioactive, or dye molecules. Assays which amplify the signals from the probe are also known; examples of which are

assays which utilize biotin and avidin, and enzyme-labeled and mediated immunoassays, such as ELISA assays.

Kits suitable for immunodiagnosis and containing the appropriate labeled reagents are constructed by packaging the appropriate materials, including the compositions of the invention, in suitable containers, along with the remaining reagents and materials (for example, suitable buffers, salt solutions, *etc.*) required for the conduct of the assay, as well as suitable set of assay instructions.

Nucleic Acid Hybridisation

"Hybridization" refers to the association of two nucleic acid sequences to one another by hydrogen bonding. Typically, one sequence will be fixed to a solid support and the other will be free in solution. Then, the two sequences will be placed in contact with one another under conditions that favor hydrogen bonding. Factors that affect this bonding include: the type and volume of solvent; reaction temperature; time of hybridization; agitation; agents to block the non-specific attachment of the liquid phase sequence to the solid support (Denhardt's reagent or BLOTTO); concentration of the sequences; use of compounds to increase the rate of association of sequences (dextran sulfate or polyethylene glycol); and the stringency of the washing conditions following hybridization. See Sambrook *et al.* [*supra*] Volume 2, chapter 9, pages 9.47 to 9.57.

"Stringency" refers to conditions in a hybridization reaction that favor association of very similar sequences over sequences that differ. For example, the combination of temperature and salt concentration should be chosen that is approximately 120 to 200°C below the calculated T_m of the hybrid under study. The temperature and salt conditions can often be determined empirically in preliminary experiments in which samples of genomic DNA immobilized on filters are hybridized to the sequence of interest and then washed under conditions of different stringencies. See Sambrook *et al.* at page 9.50.

Variables to consider when performing, for example, a Southern blot are (1) the complexity of the DNA being blotted and (2) the homology between the probe and the sequences being detected. The total amount of the fragment(s) to be studied can vary a magnitude of 10, from 0.1 to 1 μ g for a plasmid or phage digest to 10^{-9} to 10^{-8} g for a single copy gene in a highly complex eukaryotic genome. For lower complexity polynucleotides, substantially shorter blotting, hybridization, and exposure times, a smaller amount of starting polynucleotides, and lower specific activity of probes can be used. For example, a single-copy

yeast gene can be detected with an exposure time of only 1 hour starting with 1 µg of yeast DNA, blotting for two hours, and hybridizing for 4-8 hours with a probe of 10^8 cpm/µg. For a single-copy mammalian gene a conservative approach would start with 10 µg of DNA, blot overnight, and hybridize overnight in the presence of 10% dextran sulfate using a probe of greater than 10^8 cpm/µg, resulting in an exposure time of ~24 hours.

Several factors can affect the melting temperature (T_m) of a DNA-DNA hybrid between the probe and the fragment of interest, and consequently, the appropriate conditions for hybridization and washing. In many cases the probe is not 100% homologous to the fragment. Other commonly encountered variables include the length and total G+C content of the hybridizing sequences and the ionic strength and formamide content of the hybridization buffer. The effects of all of these factors can be approximated by a single equation:

$$T_m = 81 + 16.6(\log_{10} C_i) + 0.4[\%(G + C)] - 0.6(\%\text{formamide}) - 600/n - 1.5(\%\text{mismatch}).$$

where C_i is the salt concentration (monovalent ions) and n is the length of the hybrid in base pairs (slightly modified from Meinkoth & Wahl (1984) *Anal. Biochem.* 138: 267-284).

In designing a hybridization experiment, some factors affecting nucleic acid hybridization can be conveniently altered. The temperature of the hybridization and washes and the salt concentration during the washes are the simplest to adjust. As the temperature of the hybridization increases (*ie.* stringency), it becomes less likely for hybridization to occur between strands that are nonhomologous, and as a result, background decreases. If the radiolabeled probe is not completely homologous with the immobilized fragment (as is frequently the case in gene family and interspecies hybridization experiments), the hybridization temperature must be reduced, and background will increase. The temperature of the washes affects the intensity of the hybridizing band and the degree of background in a similar manner. The stringency of the washes is also increased with decreasing salt concentrations.

In general, convenient hybridization temperatures in the presence of 50% formamide are 42°C for a probe with is 95% to 100% homologous to the target fragment, 37°C for 90% to 95% homology, and 32°C for 85% to 90% homology. For lower homologies, formamide content should be lowered and temperature adjusted accordingly, using the equation above. If the homology between the probe and the target fragment are not known, the simplest approach is to start with both hybridization and wash conditions which are nonstringent. If non-specific bands or high background are observed after autoradiography, the filter can be

washed at high stringency and reexposed. If the time required for exposure makes this approach impractical, several hybridization and/or washing stringencies should be tested in parallel.

Nucleic Acid Probe Assays

Methods such as PCR, branched DNA probe assays, or blotting techniques utilizing nucleic acid probes according to the invention can determine the presence of cDNA or mRNA. A probe is said to "hybridize" with a sequence of the invention if it can form a duplex or double stranded complex, which is stable enough to be detected.

The nucleic acid probes will hybridize to the Neisserial nucleotide sequences of the invention (including both sense and antisense strands). Though many different nucleotide sequences will encode the amino acid sequence, the native Neisserial sequence is preferred because it is the actual sequence present in cells. mRNA represents a coding sequence and so a probe should be complementary to the coding sequence; single-stranded cDNA is complementary to mRNA, and so a cDNA probe should be complementary to the non-coding sequence.

The probe sequence need not be identical to the Neisserial sequence (or its complement) — some variation in the sequence and length can lead to increased assay sensitivity if the nucleic acid probe can form a duplex with target nucleotides, which can be detected. Also, the nucleic acid probe can include additional nucleotides to stabilize the formed duplex. Additional Neisserial sequence may also be helpful as a label to detect the formed duplex. For example, a non-complementary nucleotide sequence may be attached to the 5' end of the probe, with the remainder of the probe sequence being complementary to a Neisserial sequence. Alternatively, non-complementary bases or longer sequences can be interspersed into the probe, provided that the probe sequence has sufficient complementarity with the a Neisserial sequence in order to hybridize therewith and thereby form a duplex which can be detected.

The exact length and sequence of the probe will depend on the hybridization conditions, such as temperature, salt condition and the like. For example, for diagnostic applications, depending on the complexity of the analyte sequence, the nucleic acid probe typically contains at least 10-20 nucleotides, preferably 15-25, and more preferably at least 30 nucleotides, although it may be shorter than this. Short primers generally require cooler temperatures to form sufficiently stable hybrid complexes with the template.

Probes may be produced by synthetic procedures, such as the triester method of Matteucci *et al.* [*J. Am. Chem. Soc.* (1981) 103:3185], or according to Urdea *et al.* [*Proc. Natl. Acad. Sci. USA* (1983) 80: 7461], or using commercially available automated oligonucleotide synthesizers.

The chemical nature of the probe can be selected according to preference. For certain applications, DNA or RNA are appropriate. For other applications, modifications may be incorporated *eg.* backbone modifications, such as phosphorothioates or methylphosphonates, can be used to increase *in vivo* half-life, alter RNA affinity, increase nuclease resistance *etc.* [*eg.* see Agrawal & Iyer (1995) *Curr Opin Biotechnol* 6:12-19; Agrawal (1996) *TIBTECH* 14:376-387]; analogues such as peptide nucleic acids may also be used [*eg.* see Corey (1997) *TIBTECH* 15:224-229; Buchardt *et al.* (1993) *TIBTECH* 11:384-386].

One example of a nucleotide hybridization assay is described by Urdea *et al.* in international patent application WO92/02526 [see also US patent 5,124,246].

Alternatively, the polymerase chain reaction (PCR) is another well-known means for detecting small amounts of target nucleic acids. The assay is described in: Mullis *et al.* [*Meth. Enzymol.* (1987) 155: 335-350]; US patent 4,683,195; and US patent 4,683,202. Two "primer" nucleotides hybridize with the target nucleic acids and are used to prime the reaction. The primers can comprise sequence that does not hybridize to the sequence of the amplification target (or its complement) to aid with duplex stability or, for example, to incorporate a convenient restriction site. Typically, such sequence will flank the desired Neisserial sequence.

A thermostable polymerase creates copies of target nucleic acids from the primers using the original target nucleic acids as a template. After a threshold amount of target nucleic acids are generated by the polymerase, they can be detected by more traditional methods, such as Southern blots. When using the Southern blot method, the labelled probe will hybridize to the Neisserial sequence (or its complement).

Also, mRNA or cDNA can be detected by traditional blotting techniques described in Sambrook *et al* [*supra*]. mRNA, or cDNA generated from mRNA using a polymerase enzyme, can be purified and separated using gel electrophoresis. The nucleic acids on the gel are then blotted onto a solid support, such as nitrocellulose. The solid support is exposed to a labelled probe and then washed to remove any unhybridized probe. Next, the duplexes

containing the labeled probe are detected. Typically, the probe is labelled with a radioactive moiety.

EXAMPLES

The examples describe nucleic acid sequences which have been identified in *N. meningitidis*, and *N. gonorrhoeae* along with their respective and putative translation products. Not all of the nucleic acid sequences are complete *ie.* they encode less than the full-length wild-type protein.

The examples are generally in the following format:

- a nucleotide sequence which has been identified in *N. meningitidis*
- the putative translation product of said *N. meningitidis* sequence
- a computer analysis of said translation product based on database comparisons
- a corresponding nucleotide sequence identified from *N. gonorrhoeae*
- the putative translation product of said *N. gonorrhoeae* sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a corresponding nucleotide sequence identified from strain A of *N. meningitidis*
- the putative translation product of said *N. meningitidis* strain A sequence
- a comparison of the percentage of identity between the translation product of the *N. meningitidis* sequence and the *N. gonorrhoeae* sequence.
- a description of the characteristics of the protein which indicates that it might be suitably antigenic or immunogenic.

Sequence comparisons were performed at NCBI (<http://www.ncbi.nlm.nih.gov>) using the algorithms BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx, & tBLASTx [eg. see also Altschul *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* 25:2289-3402]. Searches were performed against the following databases: non-redundant GenBank+EMBL+DDBJ+PDB sequences and non-redundant GenBank CDS translations+PDB+SwissProt+SPupdate+PIR sequences.

Dots within nucleotide sequences represent nucleotides which have been arbitrarily introduced in order to maintain a reading frame. In the same way, double-underlined nucleotides were removed. Lower case letters represent ambiguities which arose during

alignment of independent sequencing reactions (some of the nucleotide sequences in the examples are derived from combining the results of two or more experiments).

Nucleotide sequences were scanned in all six reading frames to predict the presence of hydrophobic domains using an algorithm based on the statistical studies of Esposti *et al.* [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. These domains represent potential transmembrane regions or hydrophobic leader sequences.

Open reading frames were predicted from fragmented nucleotide sequences using the program ORFFINDER (NCBI).

Underlined amino acid sequences indicate possible transmembrane domains or leader sequences in the ORFs, as predicted by the PSORT algorithm (<http://www.psort.nibb.ac.jp>). Functional domains were also predicted using the MOTIFS program (GCG Wisconsin & PROSITE).

For each of the following examples: based on the presence of a putative leader sequence and/or several putative transmembrane domains (single-underlined) in the gonococcal protein, it is predicted that the proteins from *N. meningitidis* and *N. gonorrhoeae*, and their respective epitopes, could be useful antigens or immunogenic compositions for vaccines or diagnostics.

The standard techniques and procedures which may be employed in order to perform the invention (*e.g.* to utilize the disclosed sequences for vaccination or diagnostic purposes) were summarized above. This summary is not a limitation on the invention but, rather, gives examples that may be used, but are not required.

In particular, the following methods were used to express, purify and biochemically characterize the proteins of the invention.

Chromosomal DNA Preparation

N.meningitidis strain 2996 was grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20%(w/v) Sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl₃/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation.

The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading the OD at 260 nm.

Oligonucleotide design

Synthetic oligonucleotide primers were designed on the basis of the coding sequence of each ORF, using (a) the meningococcus B sequence when available, or (b) the gonococcus/meningococcus A sequence, adapted to the codon preference usage of meningococcus as necessary. Any predicted signal peptides were omitted, by designing the 5' primers to sequence immediately downstream from the predicted leader sequence.

For most ORFs, the 5' primers included two restriction enzyme recognition sites (*Bam*HI-*Nde*I, *Bam*HI-*Nhe*I, *Eco*RI-*Nde*I or *Eco*RI-*Nhe*I), depending on the restriction pattern of the gene of interest. The 3' primers included a *Xho*I or a *Hind*III restriction site (table 1). This procedure was established in order to direct the cloning of each amplification product (corresponding to each ORF) into two different expression systems: pGEX-KG (using *Bam*HI-*Xho*I, *Bam*HI-*Hind*III, *Eco*RI-*Xho*I or *Eco*RI-*Hind*III), and pET21b+ (using *Nde*I-*Xho*I, *Nhe*I-*Xho*I, *Nde*I-*Hind*III or *Nhe*I-*Hind*III).

5'-end primer tail:	<u>CGCGGATCCCATATG</u>	(<i>Bam</i> HI- <i>Nde</i> I)
	<u>CGCGGATCCGCTAGC</u>	(<i>Bam</i> HI- <i>Nhe</i> I)
	<u>CCGGAATTCTACATATG</u>	(<i>Eco</i> RI- <i>Nde</i> I)
	<u>CCGGAATTCTAGCTAGC</u>	(<i>Eco</i> RI- <i>Nhe</i> I)
3'-end primer tail:	<u>CCCGCTCGAG</u>	(<i>Xho</i> I)
	<u>CCCGCTCGAG</u>	(<i>Hind</i> III)

For cloning ORFs into the pGEX-His vector, the 5' and 3' primers contained only one restriction enzyme site (*Eco*RI, *Kpn*I or *Sal*I for the 5' primers and *Pst*I, *Xba*I, *Sph*I or *Sal*I for the 3' primers). Again restriction sites were chosen according to the particular restriction pattern of the gene (table 1).

5'-end primer tail:	(AAA) <u>AAAGAATTC</u>	(<i>Eco</i> RI)
	(AAA) <u>AAAGGTACC</u>	(<i>Kpn</i> I)
3'-end primer tail:	(AAA) <u>AAACTGCAG</u>	(<i>Pst</i> I)
	(AAA) <u>AAATCTAGA</u>	(<i>Xba</i> I)

5' or 3'-end primer tail: AAAGCATGC (*Sph*I) AAAAAAGTCGAC (*Sal*I)

As well as containing the restriction enzyme recognition sequences, the primers included nucleotides which hybridized to the sequence to be amplified. The melting temperature depended on the number and type of hybridising nucleotides in the whole primer, and was determined for each primer using the formulae:

$$T_m = 4 (G+C) + 2 (A+T) \quad (\text{tail excluded})$$

$$T_m = 64.9 + 0.41 (\% \text{ GC}) - 600/N \quad (\text{whole primer})$$

The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-55°C for the hybridising region alone.

Table 1 shows the forward and reverse primers used for each amplification. In certain cases, the sequence of the primer does not exactly match the sequence of the predicted ORF. This is because when initial amplifications were performed, the complete 5' and/or 3' sequences for some meningococcal B ORFs were not known. However the corresponding sequences had been identified in *Gonococcus* or in *Meningococcus A*. Hence, when the *Meningococcus B* sequence was incomplete or uncertain, *Gonococcal* or *Meningococcal A* sequences were used as the basis for primer design. These sequences were altered to take account of codon preference. It can be appreciated that, once the complete sequence is identified, this approach will no longer be necessary.

Oligonucleotides were synthesized using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH₄OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in either 100µl or 1.0ml of water. The OD₂₆₀ was determined using a Perkin Elmer Lambda Bio spectrophotometer and the concentration adjusted to 2-10pmol/µl.

Amplification

The standard PCR protocol was as follows: 50-200ng of genomic DNA was used as a template in the presence of 20-40µM of each oligonucleotide primer, 400-800µM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl₂), 2.5 units *TaqI* DNA polymerase (using

Perkin-Elmer AmpliTaq, GIBCO Platinum, Pwo DNA polymerase, or Tahara Shuzo Taq polymerase). In some cases, PCR was optimised by the addition of 10µl DMSO or 50µl 2M Betaine.

After a hot start (adding the polymerase during a preliminary 3 minute incubation of the whole mix at 95°C), each sample underwent a two-step amplification. The first 5 cycles were performed using the hybridization temperature that excluded the restriction enzyme tail of the primer (see above). This was followed by 30 cycles using the hybridization temperature calculated for the whole length oligos. The cycles were completed with a 10 minute extension step at 72°C. The standard cycles were as follows:

	Denaturation	Hybridisation	Elongation
First 5 cycles	30 seconds 95°C	30 seconds 50-55°C	30-60 seconds 72°C
Last 30 cycles	30 seconds 95°C	30 seconds 65-70°C	30-60 seconds 72°C

Elongation times varied according to the length of the ORF to be amplified. Amplifications were performed using either a 9600 or a 2400 Perkin Elmer GeneAmp PCR System. To check the results, 1/10 of the amplification volume was loaded onto a 1-1.5% (w/v) agarose gel and the size of each amplified fragment compared with a DNA molecular weight marker.

The amplified DNA was either loaded directly on a 1% agarose gel or first precipitated with ethanol and resuspended in a volume suitable to be loaded on a 1.0% agarose gel. The DNA fragment corresponding to the band of correct size was purified using the Qiagen Gel Extraction Kit, following the manufacturer's protocol. DNA fragments were eluted in a volume of 30µl or 50µl with either H₂O or 10mM Tris, pH 8.5.

Digestion of PCR fragments

The purified DNA corresponding to the amplified fragment was doubly-digested with the appropriate restriction enzymes for; cloning into pET-21b+ and expressing the protein as a C-terminus His-tagged fusion, for cloning into pGEX-KG and expressing the protein as a N-

ratio of 3:1 fragment/vector was ligated using 0.5 µl of NEB T4 DNA ligase (400 units/µl), in the presence of the buffer supplied by the manufacturer. The reaction was incubated at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit", following the manufacturer's instructions.

In order to introduce the recombinant plasmid in a suitable strain, 100 µl *E. coli* DH5 competent cells were incubated with the ligase reaction solution for 40 minutes on ice, then at 37°C for 3 minutes, then, after adding 800 µl LB broth, again at 37°C for 20 minutes. The cells were then centrifuged at maximum speed in an Eppendorf microfuge and resuspended in approximately 200 µl of the supernatant. The suspension was then plated on LB ampicillin (100 mg/ml).

The screening of the recombinant clones was performed by growing 5 randomly-chosen colonies overnight at 37 °C in either 2 ml (pGEX or pTC clones) or 5ml (pET clones) LB broth + 100 µg/ml ampicillin. The cells were then pelleted and the DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions, to a final volume of 30 µl. 5 µl of each individual miniprep (approximately 1g) were digested with either *NdeI/XhoI* or *BamHI/XhoI* and the whole digestion loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1Kb DNA Ladder, GIBCO). The screening of the positive clones was made on the base of the correct insert size.

For other ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated into both pET21b+ and pGEX-KG. A molar ratio of of 3:1 fragment/vector was used in a final volume of 20µl, that included 0.5µl T4 DNA ligase (400 units/µl, NEB) and ligation buffer supplied by the manufacturer. The reaction was performed at room temperature for 3 hours. In some experiments, ligation was performed using the Boheringer "Rapid Ligation Kit" and the manufacturer's protocol.

Recombinant plasmid was transformed into 100µl of competent *E. coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice then at 37°C for 3 minutes. This was followed by the addition of 800µl LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200µl of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

Screening for recombinant clones was performed by growing 5 randomly selected colonies overnight at 37°C in either 2.0ml (pGEX-KG clones) or 5.0ml (pET clones) LB broth + 100µg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1µg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

ORFs were cloned into PGEX-His, by doubly-digesting the PCR product and ligating into similarly digested vector. After cloning, recombinant plasmids were transformed into the *E.coli* host W3110. Individual clones were grown overnight at 37°C in LB broth with 50µg/ml ampicillin.

Certain ORFs may be cloned into the pGEX-HIS vector using *EcoRI-PstI* cloning sites, or *EcoRI-SalI*, or *SalI-PstI*. After cloning, the recombinant plasmids may be introduced in the *E.coli* host W3110.

Expression

Each ORF cloned into the expression vector may then be transformed into the strain suitable for expression of the recombinant protein product. 1 µl of each construct was used to transform 30 µl of *E.coli* BL21 (pGEX vector), *E.coli* TOP 10 (pTRC vector) or *E.coli* BL21-DE3 (pET vector), as described above. In the case of the pGEX-His vector, the same *E.coli* strain (W3110) was used for initial cloning and expression. Single recombinant colonies were inoculated into 2ml LB+Amp (100 µg/ml), incubated at 37°C overnight, then diluted 1:30 in 20 ml of LB+Amp (100 µg/ml) in 100 ml flasks, making sure that the OD₆₀₀ ranged between 0.1 and 0.15. The flasks were incubated at 30°C into gyratory water bath shakers until OD indicated exponential growth suitable for induction of expression (0.4-0.8 OD for pET and pTRC vectors; 0.8-1 OD for pGEX and pGEX-His vectors). For the pET, pTRC and pGEX-His vectors, the protein expression was induced by addition of 1mM IPTG, whereas in the case of pGEX system the final concentration of IPTG was 0.2 mM. After 3 hours incubation at 30°C, the final concentration of the sample was checked by OD. In order to check expression, 1ml of each sample was removed, centrifuged in a microfuge, the pellet

resuspended in PBS, and analysed by 12% SDS-PAGE with Coomassie Blue staining. The whole sample was centrifuged at 6000g and the pellet resuspended in PBS for further use.

GST-fusion proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on LB+Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600 ml of fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD₅₅₀ 0.8-1. Protein expression was induced with 0.2mM IPTG followed by three hours incubation. The culture was centrifuged at 8000 rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml cold PBS. The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia) (previously washed with PBS) and incubated at room temperature for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold PBS for 10 minutes, resuspended in 1ml cold PBS, and loaded on a disposable column. The resin was washed twice with 2ml cold PBS until the flow-through reached OD₂₈₀ of 0.02-0.06. The GST-fusion protein was eluted by addition of 700µl cold Glutathione elution buffer 10mM reduced glutathione, 50mM Tris-HCl) and fractions collected until the OD₂₈₀ was 0.1. 21µl of each fraction were loaded on a 12% SDS gel using either Biorad SDS-PAGE Molecular weight standard broad range (M1) (200, 116.25, 97.4, 66.2, 45, 31, 21.5, 14.4, 6.5 kDa) or Amersham Rainbow Marker (M") (220, 66, 46, 30, 21.5, 14.3 kDa) as standards. As the MW of GST is 26kDa, this value must be added to the MW of each GST-fusion protein.

For other ORFs, for each clone to be purified as a GST-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp. (100µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Recombinant protein expression was induced by addition of IPTG (final concentration 0.2mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

terminus GST-fusion, and for cloning into pGEX-His and expressing the protein as a N-terminus GST-His tagged fusion.

Each purified DNA fragment was incubated at 37°C for 3 hours to overnight with 20 units of appropriate restriction enzyme (New England Biolabs) in a volume of either 30 or 40 µl in the presence of suitable digestion buffer. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted in a volume of 30 µl or 50 µl with either H₂O or 10mM Tris, pH 8.5. The DNA concentration was determined by quantitative agarose gel electrophoresis (1.0% gel) in the presence of a titrated molecular weight marker.

Digestion of the cloning vectors (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG, and pGEX-His)

The vector pGEX-His is a modified pGEX-2T vector carrying a region encoding six histidine residues upstream of the thrombin cleavage site and containing the multiple cloning site of the vector pTRC99 (Pharmacia). 10 µg plasmid was double-digested with 50 units of each restriction enzyme in 200 µl reaction volume in the presence of appropriate buffer by overnight incubation at 37°C. After loading the whole digestion on a 1% agarose gel, the band corresponding to the digested vector was purified from the gel using the Qiagen QIAquick Gel Extraction Kit and the DNA was eluted in 50 µl of 10 mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD₂₆₀ of the sample, and adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

10 µg of plasmid vector was doubly-digested with 50 units of each restriction enzyme in a volume of 200 µl with the appropriate buffer overnight at 37°C. The digest was loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit. DNA was eluted in 50 µl of 10mM Tris-HCl, pH 8.5. The DNA concentration was evaluated by measuring OD_{260nm} and the concentration adjusted to 50 µg/µl. 1 µl of plasmid was used for each cloning procedure.

Cloning

For some ORFs, the fragments corresponding to each ORF, previously digested and purified, were ligated in both pET22b and pGEX-KG. In a final volume of 20 µl, a molar

The bacterial pellet was resuspended in 7.5ml cold PBS. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and mixed with 150µl Glutathione-Sepharose 4B resin (Pharmacia), previously equilibrated with PBS, and incubated at room temperature with gentle agitation for 30 min. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batchwise) with 10ml cold PBS for 10 min, resuspended in 1ml cold PBS, and loaded onto a disposable column. The resin continued to be washed with cold PBS, until the OD_{280nm} of the flow-through reached 0.02-0.01. The GST-fusion protein was eluted by addition of 700µl cold glutathione elution buffer (10mM reduced glutathione, 50mM Tris-HCl pH 8.0) and fractions collected, until the OD_{280nm} of the eluate indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. The molecular mass of the purified proteins was determined using either the Bio-Rad broad range molecular weight standard (M1) (200, 116, 97.4, 66.2, 45.0, 31.0, 21.5, 14.4, 6.5 kDa) or the Amersham Rainbow Marker (M2) (220, 66.2, 46.0, 30.0, 21.5, 14.3 kDa). The molecular weights of GST-fusion proteins are a combination of the 26 kDa GST protein and its fusion partner. Protein concentrations were estimated using the Bradford assay.

His-fusion soluble proteins large-scale purification.

For some ORFs, a single colony was grown overnight at 37°C on a LB + Amp agar plate. The bacteria were inoculated into 20ml of LB+Amp liquid culture and incubated overnight in a water bath shaker. Bacteria were diluted 1:30 into 600ml fresh medium and allowed to grow at the optimal temperature (20-37°C) to OD₅₅₀ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000 rpm at 4°C, the supernatant was discarded and the bacterial pellet was resuspended in 7.5ml cold 10mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed two times and centrifuged again. The supernatant was collected and mixed with 150µl Ni²⁺-resin (Pharmacia) (previously washed with 10mM imidazole buffer) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700g for 5 minutes at 4°C. The resin was washed twice with 10 ml cold 10mM imidazole buffer for 10 minutes,

resuspended in 1ml cold 10mM imidazole buffer and loaded on a disposable column. The resin was washed at 4°C with 2ml cold 10mM imidazole buffer until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The resin was washed with 2ml cold 20mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 20 mM imidazole, pH 8) until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl cold 250mM imidazole buffer (300 mM NaCl, 50 mM phosphate buffer, 250 mM imidazole, pH 8) and fractions collected until the O.D₂₈₀ was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

His-fusion insoluble proteins large-scale purification.

A single colony was grown overnight at 37 °C on a LB + Amp agar plate. The bacteria were inoculated into 20 ml of LB+Amp liquid culture in a water bath shaker and grown overnight. Bacteria were diluted 1:30 into 600ml fresh medium and let to grow at the optimal temperature (37°C) to O.D₅₅₀ 0.6-0.8. Protein expression was induced by addition of 1 mM IPTG and the culture further incubated for three hours. The culture was centrifuged at 8000rpm at 4°C. The supernatant was discarded and the bacterial pellet was resuspended in 7.5 ml buffer B (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8). The cells were disrupted by sonication on ice for 30 sec at 40W using a Branson sonifier B-15, frozen and thawed twice and centrifuged again. The supernatant was stored at -20°C, while the pellets were resuspended in 2 ml guanidine buffer (6M guanidine hydrochloride, 100mM phosphate buffer, 10 mM Tris-HCl, pH 7.5) and treated in a homogenizer for 10 cycles. The product was centrifuged at 13000 rpm for 40 minutes. The supernatant was mixed with 150µl Ni²⁺-resin (Pharmacia) (previously washed with buffer B) and incubated at room temperature with gentle agitation for 30 minutes. The sample was centrifuged at 700 g for 5 minutes at 4°C. The resin was washed twice with 10 ml buffer B for 10 minutes, resuspended in 1ml buffer B, and loaded on a disposable column. The resin was washed at room temperature with 2ml buffer B until the flow-through reached the OD₂₈₀ of 0.02-0.06. The resin was washed with 2ml buffer C (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the flow-through reached the O.D₂₈₀ of 0.02-0.06. The His-fusion protein was eluted by addition of 700µl elution buffer (urea 8M, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions collected until the OD₂₈₀ was 0.1. 21µl of each fraction were loaded on a 12% SDS gel.

Purification of His-fusion proteins.

For each clone to be purified as a His-fusion, a single colony was streaked out and grown overnight at 37°C on a LB/Amp (100 µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 600ml LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (20-37°C) until the OD_{550nm} reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000xg for 15 min at 4°C.

The bacterial pellet was resuspended in 7.5ml of either (i) cold buffer A (300mM NaCl, 50mM phosphate buffer, 10mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 8.8) for insoluble proteins. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6M guanidine hydrochloride, 100mM phosphate buffer, 10mM Tris-HCl, pH 7.5) and treated with a Dounce homogenizer for 10 cycles. The homogenate was centrifuged at 13 000xg for 40 min and the supernatant retained.

Supernatants for both soluble and insoluble preparations were mixed with 150µl Ni²⁺-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to manufacturers protocol. The batch-wise preparation was centrifuged at 700xg for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD_{280nm} of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (8M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 6.3) until the the OD_{280nm} of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700µl of either (i) cold elution buffer A (300mM NaCl, 50mM phosphate buffer, 250mM imidazole, pH 8.0) or (ii) elution buffer B (8 M urea, 10mM Tris-HCl, 100mM phosphate buffer, pH 4.5) and fractions

collected until the $O.D_{280nm}$ indicated all the recombinant protein was obtained. 20 μ l aliquots of each elution fraction were analyzed by SDS-PAGE using a 12% gel. Protein concentrations were estimated using the Bradford assay.

His-fusion proteins renaturation

In the cases where denaturation was required to solubilize proteins, a renaturation step was employed prior to immunization. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10%(v/v). The proteins were diluted to 200 μ g/ml using dialysis buffer I (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% (v/v) glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Alternatively, 10% glycerol was added to the denatured proteins. The proteins were then diluted to 20 μ g/ml using dialysis buffer I (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, 2M urea, pH 8.8) and dialysed against the same buffer at 4°C for 12-14 hours. The protein was further dialysed against dialysis buffer II (10% glycerol, 0.5M arginine, 50mM phosphate buffer, 5mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C.

Protein concentration was evaluated using the formula:

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

Purification of proteins

To analyse the solubility, pellets obtained from 3.0ml cultures were resuspended in 500 μ l buffer M1 (PBS pH 7.2). 25 μ l of lysozyme (10mg/ml) was added and the bacteria incubated for 15 min at 4°C. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13 000xg for 30 min at 4°C. The supernatant was collected and the pellet resuspended in buffer M2 [8M urea, 0.5M NaCl, 20mM imidazole and 0.1M NaH_2PO_4] and incubated for 3 to 4 hours at 4°C. After centrifugation, the supernatant was collected and the pellet resuspended in buffer M3 [6M guanidinium-HCl, 0.5M NaCl, 20mM imidazole and 0.1M NaH_2PO_4] overnight at 4°C. The

supernatants from all steps were analysed by SDS-PAGE. Some proteins were found to be soluble in PBS, others need urea or guanidium-HCl for solubilization.

For preparative scale purifications, 500ml cultures were induced and fusion proteins solubilized in either buffer M1, M2 or M3 using the procedure described above. Crude extracts were loaded onto a Ni-NTA superflow column (Quiagen) equilibrated with buffer M1, M2 or M3 depending on the solubilization buffer employed. Unbound material was eluted by washing the column with the same buffer. The recombinant fusion protein was eluted with the corresponding buffer containing 500mM imidazole then dialysed against the same buffer in the absence of imidazole.

Mice immunisations

20µg of each purified protein are used to immunise mice intraperitoneally. In the case of some ORFs, Balb-C mice were immunised with Al(OH)₃ as adjuvant on days 1, 21 and 42, and immune response was monitored in samples taken on day 56. For other ORFs, CD1 mice could be immunised using the same protocol. For ORFs 25 and 40, CD1 mice were immunised using Freund's adjuvant, and the same immunisation protocol was used, except that the immune response was measured on day 42, rather than 56. Similarly, for still other ORFs, CD1 mice were immunised with Freund's adjuvant, but the immune response was measured on day 49. Alternatively, 20µg of each purified protein was mixed with Freund's adjuvant and used to immunise CD1 mice intraperitoneally. For many of the proteins, the immunization was performed on days 1, 21 and 35, and immune response was monitored in samples taken on days 34 and 49. For some proteins, the third immunization was performed on day 28, rather than 35, and the immune response was measured on days 20 and 42, rather than 34 and 49.

ELISA assay (sera analysis)

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 7ml of Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10000 rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated

for 2 hours at room temperature and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200 µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200 µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN₃ in PBS) were added to each well and the plates incubated for 90 minutes at 37°C. Wells were washed three times with PBT. 100 µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 µl of substrate buffer for HRP (25 ml of citrate buffer pH5, 10 mg of O-phenildiamine and 10 µl of H₂O) were added to each well and the plates were left at room temperature for 20 minutes. 100 µl H₂SO₄ was added to each well and OD₄₉₀ was followed. The ELISA was considered positive when OD₄₉₀ was 2.5 times the respective pre-immune sera.

Alternatively, The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% Glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.3-0.4. The culture was centrifuged for 10 minutes at 10 000rpm. The supernatant was discarded and bacteria were washed once with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200µl of saturation buffer (2.7% Polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1% Tween-20, 0.1% NaN₃ in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-

phenildiamine and 10 μ l of H₂O₂) were added to each well and the plates were left at room temperature for 20 minutes. 100 μ l of 12.5% H₂SO₄ was added to each well and OD₄₉₀ was followed. The ELISA titers were calculated arbitrarily as the dilution of sera which gave an OD₄₉₀ value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD₄₉₀ of 0.4 was higher than 1:400.

FACScan bacteria Binding Assay procedure.

The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD₆₂₀. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN₃) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD₆₂₀ of 0.07. 100 μ l bacterial cells were added to each well of a Costar 96 well plate. 100 μ l of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200 μ l/well of blocking buffer in each well. 100 μ l of R-Phicoerytrin conjugated F(ab)₂ goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200 μ l/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200 μ l/well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

OMV preparations

Bacteria were grown overnight on 5 GC plates, harvested with a loop and resuspended in 10 ml 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes and the bacteria disrupted by sonication for 10' on ice (50% duty cycle, 50% output). Unbroken cells were removed by centrifugation at 5000g for 10 minutes and the total cell envelope

fraction recovered by centrifugation at 50000g at 4°C for 75 minutes. To extract cytoplasmic membrane proteins from the crude outer membranes, the whole fraction was resuspended in 2% sarkosyl (Sigma) and incubated at room temperature for 20 minutes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, and the supernatant further ultracentrifuged at 50000g for 75 minutes to pellet the outer membranes. The outer membranes were resuspended in 10mM Tris-HCl, pH8 and the protein concentration measured by the Bio-Rad Protein assay, using BSA as a standard.

Whole Extracts preparation

Bacteria were grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30' minutes.

Western blotting

Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3 % Tris base, 1.44 % glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

Bactericidal assay

MC58 and 2996 strains were grown overnight at 37°C on chocolate agar plates. 5-7 colonies were collected and used to inoculate 7ml Mueller-Hinton broth. The suspension was incubated at 37°C on a nutator and let to grow until OD₆₂₀ was in between 0.5-0.8. The culture was aliquoted into sterile 1.5ml Eppendorf tubes and centrifuged for 20 minutes at maximum speed in a microfuge. The pellet was washed once in Gey's buffer (Gibco) and resuspended in the same buffer to an OD₆₂₀ of 0.5, diluted 1:20000 in Gey's buffer and stored at 25°C.

50µl of Gey's buffer/1% BSA was added to each well of a 96-well tissue culture plate. 25µl of diluted (1:100) mice sera (dilution buffer: Gey's buffer/0.2% BSA) were added to each well and the plate incubated at 4°C. 25µl of the previously described bacterial suspension were added to each well. 25µl of either heat-inactivated (56°C waterbath for 30 minutes) or normal baby rabbit complement were added to each well. Immediately after the addition of the baby rabbit complement, 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 0). The 96-well plate was incubated for 1 hour at 37°C with rotation and then 22µl of each sample/well were plated on Mueller-Hinton agar plates (time 1). After overnight incubation the colonies corresponding to time 0 and time 1h were counted.

Gene Variability

The *ORF4* and *919* genes were amplified by PCR on chromosomal DNA extracted from various *Neisseria* strains (see list of strains). The following oligonucleotides used as PCR primers were designed in the upstream and downstream regions of the genes:

orf 4.1	(forward) CGAATCCGGACGGCAGGACTC
orf 4.3	(reverse) GGCAGGGAATGGCGGATTAAAG
919.1	(forward) AAAATGCCTCTCCACGGCTG or CTGCGCCCTGTGTTAAAATCCCCT
919.6	(reverse) CAAATAAGAAAGGAATTTTG or GGTATCGCAAACTTCGCCTTAATGCG

The PCR cycling conditions were:

1 cycle	2 min. at 94°
30 cycles	30 sec. at 94°
	30 sec. at ~ 54° or ~ 60° (in according to Tm of the primers)
	40 sec. at 72°
1 cycle	7 min. at 72°

The PCR products were purified from 1 % agarose gel and sequenced using the following primers:

orf 4.1 (forward)	CGAATCCGGACGGCAGGACTC
orf 4.2 (forward)	CGACCGCGCCTTTGGGACTG
orf 4.3 (reverse)	GGCAGGGAATGGCGGATTAAAG
orf 4.4 (reverse)	TCTTTGAGTTTGATCCAACC

- 919.1 (forward) AAAATGCCTCTCCACGGCTG or
CTGCGCCCTGTGTTAAAATCCCCT
- 919.2 (forward) ATCCTTCCGCTCGGCTGCG
- 919.3 (forward) AAAACAGCGGCACAATCGAC
- 919.4 (forward) ATAAGGGCTACCTCAAATC
- 919.5 (forward) GCGCGTGGATTATTTTGGG
- 919.6 (reverse) CAAATAAGAAAGGAATTTTG or
GGTATCGCAAACTTCGCCTTAATGCG
- 919.7 (reverse) CCCAAGGTAATGTAGTGCCG
- 919.8 (reverse) TAAAAAAAAGTTCGACAGGG
- 919.9 (reverse) CCGTCCGCCTGTCGTCGCC
- 919.10 (reverse) TCGTTCCGGCGGGTCTGGGG

All documents cited herein are incorporated by reference in their entireties.

The following Examples are presented to illustrate, not limit, the invention

EXAMPLE 1

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 1: Oligonucleotides used for PCR for Examples 2-10

ORF	Primer	Sequence	Restriction sites
279	Forward	CGCGGATCCCATATG-TTGCCTGCAATCACGATT <SEQ ID 3021>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA <SEQ ID 3022>	XhoI
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA <SEQ ID 3023>	BamHI-NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC <SEQ ID 3024>	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT <SEQ ID 3025>	BamHI-NdeI
	Reverse	CCCGCTCGAG-ATTACTTTTTTGATGTCGAC <SEQ ID 3026>	XhoI
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC <SEQ ID 3027>	BamHI-NdeI
	Reverse	CCCGCTCGAG-CGGGCGGTATTCGGG <SEQ ID 3028>	XhoI
121	Forward	CGCGGATCCCATATG-GAAACACAGCTTTACAT <SEQ ID 3029>	BamHI-NdeI

128	Reverse	CCCGCTCGAG-ATAATAATATCCCGCGCCC <SEQ ID 3030>	XhoI
	Forward	CGCGGATCCCATATG-ACTGACAACGCACT <SEQ ID 3031>	BamHI-NdeI
206	Reverse	CCCGCTCGAG-GACCGCGTTGTCGAAA <SEQ ID 3032>	XhoI
	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA <SEQ ID 3033>	BamHI-NdeI
287	Reverse	CCCGCTCGAG-TTCTGTAAAAAAGTATGTGC <SEQ ID 3034>	XhoI
	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG <SEQ ID 3035>	EcoRI-NheI
406	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC <SEQ ID 3036>	XhoI
	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG <SEQ ID 3037>	BamHI-NdeI
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG <SEQ ID 3038>	XhoI

Localization of the ORFs

The following DNA and amino acid sequences are identified by titles of the following form: [g, m, or a] [#].[seq or pep], where "g" means a sequence from *N. gonorrhoeae*, "m" means a sequence from *N. meningitidis B*, and "a" means a sequence from *N. meningitidis A*; "#" means the number of the sequence; "seq" means a DNA sequence, and "pep" means an amino acid sequence. For example, "g001.seq" refers to an *N. gonorrhoeae* DNA sequence, number 1. The presence of the suffix "-1" to these sequences indicates an additional sequence found for the same ORF, thus, data for an ORF having both an unaffixed and a suffixed sequence designation applies to both such designated sequences. Further, open reading frames are identified as ORF #, where "#" means the number of the ORF, corresponding to the number of the sequence which encodes the ORF, and the ORF designations may be suffixed with ".ng" or ".a", indicating that the ORF corresponds to a *N. gonorrhoeae* sequence or a *N. meningitidis A* sequence, respectively. The word "partial" before a sequence indicates that the sequence may be a partial or a complete ORF. Computer analysis was performed for the comparisons that follow between "g", "m", and "a" peptide sequences; and therein the "pep" suffix is implied where not expressly stated. Further, in the event of a conflict between the text immediately preceding and describing which sequences are being compared, and the

designated sequences being compared, the designated sequence controls and is the actual sequence being compared.

ORF: contig:

279 gnm4.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3039>:

m279.seq

```

1  ATAACGCGGA TTTGCGGCTG CTTGATTCA ACGGTTTCA GGGCTTCGGC
51  AAGTTTGTGCG GCGGCGGGTT TCATCAGGCT GCAATGGGAA GGTACGGACA
101 CGGGCAGCGG CAGGGCGCGT TTGGCACCGG CTTCTTTGGC GGCAGCCATG
151 GCGCGTCCGA CGGCGGCGGC GTTGCTGCA ATCACGATTT GTCCGGGTGA
201 GTTGAAGTTG ACGGCTTCGA CCACTTCGCT TTGGGCGGCT TCGGCACAAA
251 TGGCTTTAAC CTGCTCATCT TCCAAGCCGA GAATCGCCGC CATTGCGCCC
301 ACGCCTTGCG GTACGGCGGA CTGCATCAGT TCGGCGCGCA GGCACACGAG
351 TTTGACCGCG TCGGCAAAAT TCAATGCGCC GCGGCAACG AGTGCGGTGT
401 ATTCGCCGAG GCTGTGTCCG GCAACGCGG CAGGCGTTT GCCGCCGCT
451 TCTAAATAG

```

This corresponds to the amino acid sequence <SEQ ID 3040; ORF 279>:

m279.pep

```

1  ITRICGCLIS TVFRASASLS AAGFIRLOWE GTDTGSGRAR LAPASLAAAM
51  ARPTAAALPA ITICPGELKL TASTTSLWAA SAQMALTCSS SKPRIAAIAP
101 TPCGTADCIS SARRRSLTA SAKFNAPAAT SAVYSPRLCP ATAAGVLPPA
151 SK*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3041>:

g279.seq

```

1  atgacgcgga tttgcggctg cttgatttca acggttttga gtgtttcggc
51  aagtttgtcg gcggcggtt tcatcaggct gcaatgggaa ggaacggata
101 ccggcagcgg cagggcgcgT ttggctccgg cttctttggc ggcagccatg
151 gtgcgtccga cggcgcggc gttgcctgca atcacgactt gtccgggcga
201 gttgaagttg acggcttcga ccacttcgcc ctgtgcggat tcggcacaaa
251 tctgcctgac ctgttcattt tccaaaccca aaatggccgc cattgcccct
301 acgccttgcg gtacggcgga ctgcatcagt tcggcgcgca ggcggacgag
351 tttgacggca tcggcaaaat ccaatgcttc ggcgcgaca agcgcggtgt
401 attcgccgag gctgtgtccg gcaacggcgg caggcgttt gccgccact
451 tccaaatag

```

This corresponds to the amino acid sequence <SEQ ID 3042; ORF 279.ng>:

g279.pep

```

1  MTRICGCLIS TVLSVSASLS AAGFIRLOWE GTDTGSGRAR LAPASLAAAM
51  VRPTAAALPA ITTCPGELKL TASTTSPCAD SAQICLTCSS SKPKMAAIAP
101 TPCGTADCIS SARRRSLTA SAKSNASAAT SAVYSPRLCP ATAAGVLPP
151 SK*

```

ORF 279 shows 89.5% identity over a 152 aa overlap with a predicted ORF (ORF 279.ng) from *N. gonorrhoeae*:

	10	20	30	40	50	60
m279 . pep	ITRICGCLISTVFRASASLSAAGFIRLOWEGTDTGSGRARLAPASLAAAMARPTAAALPA					
	: : : : : : :					
g279	MTRICGCLISTVLSVSASLSAAGFIRLOWEGTDTGSGRARLAPASLAAAMVRPTAAALPA					
	10	20	30	40	50	60

	70	80	90	100	110	120
m279.pep	ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA					
	: :					
g279	ITTCPGELKLTASTTSPCADSAQICLTCSSSKPKMAAIAPTPCGTADCISSARRRTSLTA					
	70	80	90	100	110	120
	130	140	150			
m279.pep	SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX					
g279	SAKSNASAATSAVYSPRLCPATAAGVLPPTSKX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3043>:

```
a279.seq
1  ATGACNCNGA TTTGCGGCTG CTTGATTTC ACGGTTTNN GGGCTTCGGC
51  GAGTTTGTG GCGGCGGGT TCATGAGGCT GCAATGGGA GGTACNGACA
101 CNGGCAGCG CAGGCGCGT TTGGCGCCG CTTCTTTGC GGCAGCATA
151 GCGCGCTCG CCGCGCGGC ATTGCCTGC ATCAGCACT GTCCGGGCGA
201 GTTGAAGTG ACGGCTTCA CACTTCATC CTGTGCGGAT TCGGCGCAA
251 TTTGTTTTAC CTGTTTCAT TCCAAGCCG GAATCGCCG CATGCGCCC
301 ACGCCTTGC GTACGGCGG CTGCATCAG TCGGCGCGC NGCGCACGAG
351 TTTGACCGC TCGGCAAAAT CCAATGCGC GCGGCAACN AGTGGGTGT
401 ATTGCGCGA GCTGTGTCC GCAACGGCG CAGGCGTTT GCCGCCGCT
451 TCCGAATAG
```

This corresponds to the amino acid sequence <SEQ ID 3044; ORF 279.a>:

```
a279.pep
1  MTXICGLIS TVXRASASLS AAGFMRLQWE GTDTGSGRAR LAPASLAASI
51  ARSTAAALPA ITTCPGELKL TASTTSSCAD SAQICFTCSS SKPRIAAIAP
101 TPCGTADCI SARKRTSLTA SAKSNAPAAT SAVYSPXLCP ATAAGVLPPA
151 SE*
```

m279/a279 ORFs 279 and 279.a showed a 88.2% identity in 152 aa overlap

	10	20	30	40	50	60
m279.pep	ITRICGLISTVFRASASLSAAGFIRLQWEGTDTGSGRARLAPASLAAMARPTAAALPA					
	:					
a279	MTXICGLISTVXRASASLSAAGFMRLQWEGTDTGSGRARLAPASLAASIRSTAAALPA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m279.pep	ITICPGELKLTASTTSLWAASAQMALTCSSSKPRIAAIAPTPCGTADCISSARRRTSLTA					
	: :					
a279	ITTCPGELKLTASTTSSCADSAQICFTCSSSKPRIAAIAPTPCGTADCISSARKRTSLTA					
	70	80	90	100	110	120
	130	140	150			
m279.pep	SAKFNAPAATSAVYSPRLCPATAAGVLPPASKX					
a279	SAKSNAPAATSAVYSPXLCPATAAGVLPPASEX					
	130	140	150			

519 and 519-1 gnm7.seq

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3045>:

```
m519.seq (partial)
1  ..TCCGTTATCG GCGTATGGA GTTGGACAAA ACGTTTGAAG AACGCGACGA
51  AATCAACAGT ACTGTTGTTG CGGCTTTGGA CGAGGCGGCC GGGgCTTgGG
101 GTGTGAAGGT TTTGCGTTAT GAGATTAAAG ACTTGGTTCC GCCGCAAGAA
151 ATCCTTCGCT CAATGCAGGC GCAAATTACT GCCGAACGCG AAAAACGCGC
201 CCGTATCGCC GAATCCGAAG GTCGTAAAT CGAACAAATC AACCTTGCCA
```

```

251 GTGGTCAGCG CGAAGCCGAA ATCCAACAAT CCGAAGGCGA GGCTCAGGCT
301 GCGGTCAATG CGTCAAATGC CGAGAAAATC GCCCGCATCA ACCGCGCCAA
351 AGGTGAAGCG GAATCCTTGC GCCTTGTTGC CGAAGCCAAT GCCGAAGCCA
401 TCCGTCAAAT TGCCGCCGCC CTTCAAACCC AAGGCGGTGC GGATGCGGTC
451 AATCTGAAGA TTGCGGAACA ATACGTCGCT GCGTTCAACA ATCTTGCCAA
501 AGAAAGCAAT ACGCTGATTA TGCCCGCCAA TGTGCGGAC ATCGGCAGCC
551 TGATTCTGCG CGGTATGAAA ATTATCGACA GCAGCAAAAC CGCCAAaTAA

```

This corresponds to the amino acid sequence <SEQ ID 3046; ORF 519>:

```

m519.pep (partial)
1 ..SVIGRMELDK TFEERDEINS TVVAALDEAA GAWGVKVLRY EIKDLVPPQE
51 ILRSMQAQIT AEREKRARIA ESEGRKIEQI NLAGSQREAE IQQSEGEAQA
101 AVNASNAEKI ARINRAKGEA ESLRLVAEAN AEAIRQIAAA LQTGGGADAV
151 NLKIAEQYVA AFNNLAKESN TLIMPANVAD IGSILISAGMK IIDSSTAK*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3047>:

```

g519.seq
1 atggaatttt tcattatctt gttggcagcc gtcgccgttt tcggcttcaa
51 atcctttgtc gtcaccccc agcaggaagt ccacgttgtc gaaaggctcg
101 ggcgtttcca tcgcgccctg acggccggtt tgaatatatt gattcccttt
151 atcgaccgcg tcgcctaccg ccattcgctg aaagaaatcc ctttagacgt
201 acccagccag gtctgcatca cgcgcgataa tacgcaattg actgttgacg
251 gcatcatcta tttccaagta accgatccca aactcgcttc atacggttcg
301 agcaactaca ttatggcaat taccagctt gcccaaacga cgctgcgttc
351 cgttatcggt cgtatggagt tggacaaaac gtttgaagaa cgcgacgaaa
401 tcaacagtac cgtcgtcttc gccctcgatg aagccgccgg ggcttggggg
451 gtgaaagtcc tccgttacga aatcaaggat ttggttccgc cgcaagaaat
501 ccttcgcgca atgcaggcac aaattaccgc cgaacgcgaa aaacgcgccc
551 gtattgccga atccgaaggc cgtaaaatcg aacaaatcaa ccttgccagt
601 ggtcagcgtg aagccgaaat ccaacaatcc gaaggcgagg ctcaggctgc
651 ggtcaatgcg tccaatgccg agaaaatcgc ccgcatcaac cgcgcaaaag
701 gcgaagcgga atccctgcgc cttgttgccg aagccaatgc cgaagccaac
751 cgtcaaattg ccgccgccct tcaaacccaa agcggggcgg atgcggtcaa
801 tctgaagatt gcgggacaat acgttaccgc gttcaaaaat cttgccaaag
851 aagacaatac gcggattaag cccgccaagg ttgccgaaat cgggaaccct
901 aattttcggc ggcataaaaa attttcgcca gaagcaaaaa cggccaaata
951 a

```

This corresponds to the amino acid sequence <SEQ ID 3048; ORF 519.ng>:

```

g519.pep
1 MEFFIILLAA VAVFGKFSV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGQAQAVNA SNAEKIARIN RAKGEAESLR LVAEANA EAN
251 RQIAAALQTQ SGADAVNLKI AGQYVTAFAK LAKEDNTRIK PAKVAEIGNP
301 NFRRHEKFSP EAKTAK*

```

ORF 519 shows 87.5% identity over a 200 aa overlap with a predicted ORF (ORF 519.ng) from *N. gonorrhoeae*:

```

m519/g519

m519.pep                                10      20      30
SVIGRMELDKTFEERDEINSTVVAALDEAA
g519      YFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
           90      100     110     120     130     140

m519.pep           40      50      60      70      80      90
GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLAGSQREAE

```



```
a519.seq
  1  ATGGAATTTT  TCATTATCTT  GCTGGCAGCC  GTCGTTGTTT  TCGGCTTCAA
51  ATCCTTTTGT  GTCATCCCAC  AGCAGGAAGT  CCAGGTTGTC  GAAAGGCTCG
101 GGCCTTTTCCA TCGCGCCCTG  AGCGCCGGTT  TGAATATTTT  GATTCCCTTT
151 ATCGACCGCG  TCGCCTACCG  CCATTTCGCTG  AAAGAAATCC  CTTTAGACGT
201 ACCCAGCCAG  GTCTGCATCA  CGCGCGACAA  TACGCAGCTG  ACTGTTGACG
251 GTATCATCTA  TTTCCAAGTA  ACCGACCCCA  AACTCGCCTC  ATACGGTTTC
301 AGCAACTACA  TTTATGGCGAT  TACCCAGCTT  GCCCAAACGA  CGCTGCGTTC
351 CGTTATCGGG  CGTATGGAAT  TGGACAATAA  GTTTGAAGAA  CGCGACGAAA
401 TCAACAGCAC  CGTCGTCTCC  GCGCTCGATG  AAGCCGCGCG  AGCTTGCGGT
451 GTGAAGGTTT  TGC GTTATGA  GATTAAAGAC  TTGGTCCGCG  CGCAAGAAAT
501 CCTTCGCTCA  ATGCAGGCGC  AAATTACTGC  TGAACGCGAA  AAACGCGCCC
551 GTATCGCCGA  ATCCGAAGGT  CGTAAATCG  AACAAATCAA  CCTTGCCAGT
601 GGTCAGCGCG  AAGCCGAAGT  CCAACAATCC  GAAGGCGAGG  CTCAGGCTGC
651 GGTCAATGCG  TCAATGCCG  AGAAATATCG  CGCATCAAC  CCGCCCAAG
701 GTGAAGCGGA  ATCCTTGCGC  CTTGTTGCCG  AAGCCAATGC  CGAAGCCATC
751 CGTCAAATTG  CCGCGCCCT  TCAAACCCAA  GGCGGTGCGG  ATGCGGTCAA
801 TCTGAAGATT  CGGGAACAAT  ACGTCGCCCG  GTTCAACAA  CTGCGCAAGG
851 AAAGCAATAC  GCTGATTATG  CGGCCAATG  TTGCCGACAT  CCGGACCCCT
901 ATTTCTGCCG  GTATGAAAAT  TATCGACAGC  AGCAAAACCG  CCAAAATAA
```

```
a519.pep
  1  MEFFFIILLAA VVVFQFKSFV VIPQQEVHVV ERLGRFHRL TAGLNILIPF
51  IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE REINSTVVS ALDEAAGAWG
151 VKVLYRIEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEEAQAANA SNAEKIARIN RAKGEAESLR LVAEANAFAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*
```

```

m519.pep                                10      20      30
                                         SVIGRMELDKTFEERDEINSTVVAALDEAA
a519      YFQVTPKLASYGSSNYIMAITQLAQTTLRSVIGRMELDKTFEERDEINSTVVSALDEAA
           90      100      110      120      130      140

m519.pep      40      50      60      70      80      90
GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
|||||
a519      GAWGVKVLRYEIKDLVPPQEILRSMQAQITAEREKRARIAESEGRKIEQINLASGQREAE
           150      160      170      180      190      200

           100      110      120      130      140      150

```

```

m519.pep      IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
a519          IQQSEGEAQAAVNASNAEKIARINRAKGEAESLRLVAEANAIAIRQIAAALQTQGGADAV
                210      220      230      240      250      260

                160      170      180      190      200
m519.pep      NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSSKTAKX
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||
a519          NLKIAEQYVAAFNNLAKESNTLIMPANVADIGSLISAGMKIIDSSSKTAKX
                270      280      290      300      310

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3051>:

```

m519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGTAGCC GTCGCCGTTT TCGGTTTCAA
51  ATCCTTTGTT GTCATCCAC AACAGGAAGT CCACGTTGTC GAAAGGCTGG
101 GCGGTTTCCA TCGCGCCCTG ACGGCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGACCCCA AACTCGCCTC ATACGGTTTC
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC GTTGTGTGCG GCTTTGGACG AGGCGGCCGG GGCTTGGGGT
451 GTGAAGGTTT TGCCTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC CGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAATGCCC AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGTGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCTGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAAT TATCGACAGC AGCAAAACCG CCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 3052; ORF 519-1>:

```

m519-1.
1  MEFFIILLVA VAVFGFKSFV VIPQQEVHVV ERLGRFHRAL TAGLNILIPF
51  IDRVAYRHSI KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVA ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3053>:

```

g519-1.seq
1  ATGGAATTTT TCATTATCTT GTTGGCAGCC GTCGCCGTTT TCGGCTTCAA
51  ATCCTTTGTC GTCATCCCCC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGGTTTCCA TCGCGCCCTG ACGGCGGTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTCGCTG AAAGAAATCC CTTTAGACGT
201 ACCCAGCCAG GTCTGCATCA CGCGCGATAA TACGCAATTG ACTGTTGACG
251 GCATCATCTA TTTCCAAGTA ACCGATCCCA AACTCGCCTC ATACGGTTTC
301 AGCAACTACA TTATGGCAAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAGT TGGACAAAAC GTTTGAAGAA CGCGACGAAA
401 TCAACAGTAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG GGCTTGGGGT
451 GTGAAAGTCC TCCGTTACGA AATCAAGGAT TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCGCA ATGCAGGCAC AAATTACCGC CGAACGCGAA AAACGCGCCC
551 GTATTGCCGA ATCCGAAGGC CGTAAAATCG AACAAATCAA CCTTGCCAGT
601 GGTCAGCGTG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAATGCCC AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 CGGAAGCGGA ATCCCTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GGCGGGGCGG ATGCGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTAGCCGC GTTCAACAAT CTTGCCAAAG

```

851 AAAGCAATAC GCTGATTATG CCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GCATGAAAT TATCGACAGC AGCAAAACCG CCAATAA

This corresponds to the amino acid sequence <SEQ ID 3054; ORF 519-1.ng>:

g519-1.pep
1 MEFFIILLAA VAVFGFKSFV VIPQQEVHV V ERLGRFHRAL TAGLNILIPF
51 IDRVAYRHSL KEIPLDVPSQ VCITRDNTQL TVDGIIFYQV TDPKLASYGS
101 SNYIMAITQL AQTTLRSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRA MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGQAQAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

m519-1/g519-1 ORFs 519-1 and 519-1.ng showed a 99.0% identity in 315 aa overlap

	10	20	30	40	50	60
g519-1.pep	MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL					
m519-1	MEFFIILLVAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL					
	10	20	30	40	50	60
	70	80	90	100	110	120
g519-1.pep	KEIPLDVPSQVCITRDNTQLTVDGIIFYQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG					
m519-1	KEIPLDVPSQVCITRDNTQLTVDGIIFYQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG					
	70	80	90	100	110	120
	130	140	150	160	170	180
g519-1.pep	RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE					
m519-1	RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE					
	130	140	150	160	170	180
	190	200	210	220	230	240
g519-1.pep	KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAVNASNAEKIARINRAKGEAESLR					
m519-1	KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAVNASNAEKIARINRAKGEAESLR					
	190	200	210	220	230	240
	250	260	270	280	290	300
g519-1.pep	LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL					
m519-1	LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL					
	250	260	270	280	290	300
	310					
g519-1.pep	ISAGMKIIDSSKTAKX					
m519-1	ISAGMKIIDSSKTAKX					
	310					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3055>:

a519-1.seq
1 ATGGAATTTT TCATTATCTT GCTGGCAGCC GTCGTTGTTT TCGGCTTCAA
51 ATCCTTTGTT GTCATCCAC AGCAGGAAGT CCACGTTGTC GAAAGGCTCG
101 GCGGTTTCCA TCGCGCCCTG ACGGCCGCTT TGAATATTTT GATTCCCTTT
151 ATCGACCGCG TCGCCTACCG CCATTGCTG AAAGAAATCC CTTTAGACGT
201 ACCAGCCAG GTCTGCATCA CGCGCGACAA TACGCAGCTG ACTGTTGACG
251 GTATCATCTA TTTCCAAGTA ACCGACCCA AACTCGCTC ATACGGTTCG
301 AGCAACTACA TTATGGCGAT TACCCAGCTT GCCCAAACGA CGCTGCGTTC
351 CGTTATCGGG CGTATGGAAT TGGACAAAAC GTTTGAAGAA CGCGACGAAA

```

401 TCAACAGCAC CGTCGTCTCC GCCCTCGATG AAGCCGCCGG AGCTTGGGGT
451 GTGAAGGTTT TGCGTTATGA GATTAAAGAC TTGGTTCCGC CGCAAGAAAT
501 CCTTCGCTCA ATGCAGGCGC AAATTACTGC TGAACGCGAA AAACGCGCCC
551 GTATCGCCGA ATCCGAAGGT CGTAAATCG AACAAATCAA CCTTGCCAGT
601 GGTGAGCGCG AAGCCGAAAT CCAACAATCC GAAGGCGAGG CTCAGGCTGC
651 GGTCAATGCG TCAAATGCCG AGAAAATCGC CCGCATCAAC CGCGCCAAAG
701 GTGAAGCGGA ATCCTTGCGC CTTGTTGCCG AAGCCAATGC CGAAGCCATC
751 CGTCAAATTG CCGCCGCCCT TCAAACCCAA GCGGGTGCGG ATGCGGGTCAA
801 TCTGAAGATT GCGGAACAAT ACGTCGCCGC GTTCAACAAT CTTGCCAAAG
851 AAAGCAATAC GCTGATTATG CCCGCCAATG TTGCCGACAT CGGCAGCCTG
901 ATTTCTGCCG GTATGAAAT TATCGACAGC AGCAAACCG CCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 3056; ORF 519-1.a>:

a519-1.pep.

```

1  MEFFIILLAA VVVFQKSFV VVIPQEVHV ERLGRFHRAL TAGLNILIPF
51  IDRVAYRHS KEIPLDVPS VCITRDNTQL TVDGIIFQV TDPKLASYGS
101 SNYIMAITQL AQTTLSVIG RMELDKTFEE RDEINSTVVS ALDEAAGAWG
151 VKVLRYEIKD LVPPQEILRS MQAQITAERE KRARIAESEG RKIEQINLAS
201 GQREAEIQQS EGEAQAAVNA SNAEKIARIN RAKGEAESLR LVAEANAIAI
251 RQIAAALQTQ GGADAVNLKI AEQYVAAFNN LAKESNTLIM PANVADIGSL
301 ISAGMKIIDS SKTAK*

```

m519-1/a519-1 ORFs 519-1 and 519-1.a showed a 99.0% identity in 315 aa overlap

```

          10      20      30      40      50      60
a519-1.pep MEFFIILLAAVVVFQKSFVVVIPQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS
|||||::|:|||||
m519-1      MEFFIILLVAVVVFQKSFVVVIPQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS
          10      20      30      40      50      60

          70      80      90      100     110     120
a519-1.pep KEIPLDVPSQVCITRDNTQLTVDGIIFQVTDPKLASYGSSNYIMAITQLAQTTLSVIG
|||||
m519-1      KEIPLDVPSQVCITRDNTQLTVDGIIFQVTDPKLASYGSSNYIMAITQLAQTTLSVIG
          70      80      90      100     110     120

          130     140     150     160     170     180
a519-1.pep RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAITAERE
|||||
m519-1      RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAITAERE
          130     140     150     160     170     180

          190     200     210     220     230     240
a519-1.pep KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
|||||
m519-1      KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
          190     200     210     220     230     240

          250     260     270     280     290     300
a519-1.pep LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
|||||
m519-1      LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
          250     260     270     280     290     300

          310
a519-1.pep ISAGMKIIDSSKTAKX
|||||
m519-1      ISAGMKIIDSSKTAKX
          310

```

m576.seq.. (partial)

1	..	ATGCAGCAGG	CAAGCTATGC	GATGGGCGTG	GACATCGGAC	GCTCCCTGAA
51		GCAAAATGAAG	GAACACGGGC	CGGAAATGCA	TTTGAAAGTC	TTTACCGAAG
101		CCATGTCAGGC	AGTGTATGAC	GGCAAAGAAA	TCAAAATGAC	CGAAGAGCAG
151		GCTCAGGAAG	TCATGATGAA	ATTCTTCAG	GAACAACGAC	CTAAAGCCGT
201		AGAAAAACAC	AAGGCGGACG	CGAAGGCCAA	TAAAGAAAAA	GGCGAAGCCT
251		TTCTGAAAGA	AAATGCCGCC	AAAGACGGCG	TGAAGACCAC	TGCTTCCGGC
301		CTGCAATACA	AAATCACCAA	ACAGGGCGAA	GGCAACACAG	CGACCTAAAG
351		CGACATCGTT	ACCGTGGAAT	ACGAAGGCCG	CTGATTGAC	GGTACGSTAT
401		TCGACAGCAG	CAAAGCCAA	GGCCGCCCGG	TCACCTTCCC	TTTGAGCCAA
451		GTGATTCGCG	GTTGGACCGA	AGCGGTACAG	CTTCTGAAAG	AAGGCGCGCA
501		AGCCACGTTT	TACATCCGCT	CCAACCTTGC	TACCCGCGAA	CAGGGTCCGG
551		GCGACAAAT	CGGTCCGAAC	GCCACTTTGG	TATTTGATGT	GAAACTGGTC
601		AAAATCGGG	CACCCGAAAA	CGCGCCCGCC	AAGCAGCCGG	CTCAAGTCGA
651		CATCAAAAAA	GTAATTAA			

m576.pap.. (partial)

```

1  ..MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
51  AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENKAA KDGVKTTASG
101 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
151 VIPGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFVDKLV
201 KIGAPENAPA KOPAOGDITKK VN*

```

g576.seq.. (partial)

1	..atgggcgtgg	acatcggacg	ctccctgaaa	caaatgaagg	aacagggcgc
51	ggaaatcgat	ttgaaagtct	ttaccgatgc	catgcaggca	gtgtatgacg
101	gcaaaagaaat	caaaatgacc	gaagagcagg	ccagggaagt	gatgatgaaa
151	ttcctgcagg	agcagcaggc	taagacgcta	gaaaaacaca	aggcggatgc
201	gaaggccaac	aaagaaaaaag	gcgaagcctt	cctgaaggaa	aatgccgcgc
251	aagacggcgt	gaagaccact	gcttccggtc	tgcagtacaa	aatcacccaa
301	cagggtgaag	gcaaacagcc	gacaaaaagc	gacatcgtaa	cogtggaata
351	cgaaggccgc	ctgattgacg	gtaccgtatt	cgacagcagc	aaagcgaacg
401	gcggcccgcc	caccttccct	ttagcgcaag	tgattccggg	ttggaccgaa
451	ggcctactgc	ttctgaaaaga	aggcgcgcaa	gccacgttct	acatcccgtc
501	caaccttgcc	taccgcgaac	agggttcggg	cgaaaaaatc	ggtccgaacg
551	ccactttggt	atttgacgtg	aaactggtca	aatatcggcg	accgaaaaac
601	gcgcccgcga	agcagccgga	tcaagtcgac	atacaaaaag	taataataa

q576.pep.. (partial)

```

1  ..MGVDIGRSLK QMKEQGAeid LKVETDAMQA VYDGKEIKMT EEQAQEVMMK
51  FLQEQQAKAV EKHKADAKAN KEKGAEFLKE NAAEDGVKTT ASGLQYKITK
101 QGEGKQPTKD DIVTVEYEGR LIDGTVFSS KANGGPATFP LSQVIGWTE
151 GVRLLKEGGE ATFYIPSNLA YREQGAGEKI GPNATLVFDV KLVKIGAPEN
201 APAKOPDOVD IKKN*

```

Homology with a predicted ORF from *N. gonorrhoeae*

m576.pep

```

g576          MGVDIGRSLKQMKEQGAIEDLVFTDAMQAVYDGKEIKMTEEQAQEVMMKFLQ
                10      20      30      40      50

                70      80      90      100     110     120
m576.pep      EQQAKAVEKHKADAKANKEKGEAFLKENAAKDGVKTTASGLQYKITKQGEGKQPTKDDIV
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          EQQAKAVEKHKADAKANKEKGEAFLKENAAEDGVKTTASGLQYKITKQGEGKQPTKDDIV
                60      70      80      90      100     110

                130     140     150     160     170     180
m576.pep      TVEYEGRLIDGTVFDSSKANGGPVTFPLSQVIPGWTEGVQLLKEGGEATFYIPSNLAYRE
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          TVEYEGRLIDGTVFDSSKANGGPATFPLSQVIPGWTEGVRLLEKGEATFYIPSNLAYRE
                120     130     140     150     160     170

                190     200     210     220
m576.pep      QGAGDKIGPNATLVFDVKLVKIGAPENAPAKQPAQVDIKKVN*
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g576          QGAGEKIGPNATLVFDVKLVKIGAPENAPAKQPDQVDIKKVN*
                180     190     200     210

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3061>:

```

a576.seq
1   ATGAACACCA TTTTCAAAT CAGCGCACTG ACCCTTTCGG CCGCTTGGC
51  ACTTTCCGCC TCGCGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
101 CTGCCGCGCG TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCACG
151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
251 CCATGCAGGC AGTGTATGAC GGCAAGAGAA TCAAATGAC  CGAAGAGCAG
301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAGCCGT
351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAACAGC  CGACCAAGA
501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
551 TCGACAGCAG CAAAGCCAAC GGCGGCCCGG TCACCTTCCC TTTGAGCCAA
601 GTGATTCTGG GTTGGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGGCGA
651 AGCCACGTTT TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGCGG
701 GCGACAAAAT CGGCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
801 CATCAAAAAA GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3062; ORF 576.a>:

```

a576.pep
1   MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASSA QGDTSSIGST
51  MQQASYAMGV DIGRSLKQMK EQGAIEDLV FTEAMQAVYD GKEIKMTEEQ
101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
251 KIGAPENAPA KQPAQVDIKK VN*

```

m576/a576 ORFs 576 and 576.a showed a 99.5% identity in 222 aa overlap

```

                10      20      30
m576.pep      MQQASYAMGV DIGRSLKQMKEQGAIEDLV
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a576          CGKKEAAPASASEPAASSAQGDTSSIGSTMQQASYAMGV DIGRSLKQMKEQGAIEDLV
                30      40      50      60      70      80

                40      50      60      70      80      90
m576.pep      FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a576          FTEAMQAVYDGKEIKMTEEQAQEVMMKFLQEQQAKAVEKHKADAKANKEKGEAFLKENAA
                90      100     110     120     130     140

```

```

      100      110      120      130      140      150
m576.pep  KDGVKTTASGLQYKITKQEGEKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a576      KDGVKTTASGLQYKITKQEGEKQPTKDDIVTVEYEGRLIDGTVFDSSKANGGPVTFPLSQ
          150      160      170      180      190      200

      160      170      180      190      200      210
m576.pep  VIPGWTEGVQLLEKGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
a576      VILGWTEGVQLLEKGEATFYIPSNLAYREQGAGDKIGPNATLVFDVKLVKIGAPENAPA
          210      220      230      240      250      260

      220
m576.pep  KQPAQVDIKKVN
          ||||||||||
a576      KQPAQVDIKKVN
          270

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3063>:

```

m576-1.seq
1  ATGAACACCA  TTTTCAAAAT  CAGCGCACTG  ACCCTTTCCG  CCGCTTTGGC
51  ACTTTCCGCC  TGCGGCAAAA  AAGAAGCCGC  CCCCGCATCT  GCATCCGAAC
101 CTGCCGCCGC  TTCTTCCGCG  CAGGGCGACA  CCTCTTCGAT  CGGCAGCACG
151 ATGCAGCAGG  CAAGCTATGC  GATGGGCGTG  GACATCGGAC  GCTCCCTGAA
201 GCAAATGAAG  GAACAGGGCG  CGGAAATCGA  TTTGAAAGTC  TTTACCGAAG
251 CCATGCAGGC  AGTGTATGAC  GGCAAAGAAA  TCAAATGAC   CGAAGAGCAG
301 GCTCAGGAAG  TCATGATGAA  ATTCCTTCAG  GAACAACAGG  CTAAGCCGT
351 AGAAAAACAC  AAGGCGGACG  CGAAGGCCAA  TAAAGAAAAA  GGCGAAGCCT
401 TTCTGAAAGA  AAATGCCGCC  AAAGACGGCG  TGAAGACCAC  TGCTTCCGGC
451 CTGCAATACA  AAATCACCAA  ACAGGGCGAA  GGCAAACAGC  CGACCAAGA
501 CGACATCGTT  ACCGTGGAAT  ACGAAGGCCG  CCTGATTGAC  GGTACGGTAT
551 TCGACAGCAG  CAAAGCCAAC  GGCGGCCCGG  TCACCTTCCC  TTTGAGCCAA
601 GTGATTCCGG  GTTGGACCGA  AGGCGTACAG  CTTCTGAAAG  AAGGCGGCGA
651 AGCCACGTTC  TACATCCCGT  CCAACCTTGC  CTACCGCGAA  CAGGGTGCGG
701 GCGACAAAAT  CGGTCCGAAC  GCCACTTGG  TATTGTATGT  GAAACTGGTC
751 AAAATCGGCG  CACCCGAAAA  CGCGCCCGCC  AAGCAGCCGG  CTCAGTCTGA
801 CATCAAAAAA  GTAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 3064; ORF 576-1>:

```

m576-1.pep
1  MNTIFKISAL  TLSAALALSA  CGKKEAAPAS  ASEPAAASSA  QGDTSSIGST
51  MQQASYAMGV  DIGRSLKQMK  EQGAEIDLKV  FTEAMQAVYD  GKEIKMTEEQ
101 AQEVMMKFLQ  EQQAKAVEKH  KADAKANKEK  GEAFLENAA  KDGVKTTASG
151 LQYKITQGE  GKQPTKDDIV  TVEYEGRLID  GTVFDSSKAN  GGPVTFPLSQ
201 VIPGWTEGVQ  LLKEGGEATF  YIPSNLAYRE  QGAGDKIGPN  ATLVFDVKLV
251 KIGAPENAPA  KQPAQVDIKK  VN*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3065>:

```

g576-1.seq
1  ATGAACACCA  TTTTCAAAAT  CAGCGCACTG  ACCCTTTCCG  CCGCTTTGGC
51  ACTTTCCGCC  TGCGGCAAAA  AAGAAGCCGC  CCCCGCATCT  GCATCCGAAC
101 CTGCCGCCGC  TTCTGCCGCG  CAGGGCGACA  CCTCTTCAAT  CGGCAGCACG
151 ATGCAGCAGG  CAAGCTATGC  AATGGGCGTG  GACATCGGAC  GCTCCCTGAA
201 ACAAATGAAG  GAACAGGGCG  CGGAAATCGA  TTTGAAAGTC  TTTACCGATG
251 CCATGCAGGC  AGTGTATGAC  GGCAAAGAAA  TCAAATGAC   CGAAGAGCAG
301 GCCCAGGAAG  TGATGATGAA  ATTCCTGCAG  GAGCAGCAGG  CTAAGCCGT
351 AGAAAAACAC  AAGGCGGATG  CGAAGGCCAA  CAAAGAAAAA  GGCGAAGCCT
401 TCCTGAAGGA  AAATGCCGCC  AAAGACGGCG  TGAAGACCAC  TGCTTCCGGT
451 CTGCAGTACA  AAATCACCAA  ACAGGGTGAA  GGCAAACAGC  CGACAAAAAG
501 CGACATCGTT  ACCGTGGAAT  ACGAAGGCCG  CCTGATTGAC  GGTACCGTAT
551 TCGACAGCAG  CAAAGCCAAC  GGCGGCCCGG  CCACCTTCCC  TTTGAGCCAA
601 GTGATTCCGG  GTTGGACCGA  AGGCGTACGG  CTTCTGAAAG  AAGGCGGCGA
651 AGCCACGTTC  TACATCCCGT  CCAACCTTGC  CTACCGCGAA  CAGGGTGCGG

```

701 GCGAAAAAAT CGGTCCGAAC GCCACTTTGG TATTTGACGT GAAACTGGTC
 751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG ATCAAGTCTGA
 801 CATCAAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3066; ORF 576-1.ng>:

g576-1.pep
 1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASAA QGDTSSIGST
 51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTDAMQAVYD GKEIKMTEEQ
 101 AQEVMMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
 151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPATFPLSQ
 201 VIPGWTEGVR LLKEGGEATF YIPSNLAYRE QGAGEKIGPN ATLVFDVKLV
 251 KIGAPENAPA KQPDQVDIKK VN*

g576-1/m576-1 ORFs 576-1 and 576-1.ng showed a 97.8% identity in 272 aa overlap

	10	20	30	40	50	60
g576-1.pep	MNTIFKISAL	TL	SAALALSA	CGKKEAAPAS	ASEPAASAA	QGDTSSIGSTMQQASYAMGV
m576-1	MNTIFKISAL	TL	SAALALSA	CGKKEAAPAS	ASEPAASAA	QGDTSSIGSTMQQASYAMGV
	10	20	30	40	50	60
	70	80	90	100	110	120
g576-1.pep	DIGRSLKQMK	EQGAEIDLKV	FTDAMQAVYD	GKEIKMTEEQ	AQEVMMKFLQ	EQQAKAVEKH
m576-1	DIGRSLKQMK	EQGAEIDLKV	FTDAMQAVYD	GKEIKMTEEQ	AQEVMMKFLQ	EQQAKAVEKH
	70	80	90	100	110	120
	130	140	150	160	170	180
g576-1.pep	KADAKANKEK	GEAFLENAA	AKDGVKTTAS	GLQYKITKQGE	GKQPTKDDIV	TVEYEGRLID
m576-1	KADAKANKEK	GEAFLENAA	AKDGVKTTAS	GLQYKITKQGE	GKQPTKDDIV	TVEYEGRLID
	130	140	150	160	170	180
	190	200	210	220	230	240
g576-1.pep	GTVFDSSKAN	GGPATFPLSQ	VIPGWTEGVR	LLKEGGEATF	YIPSNLAYRE	QGAGEKIGPN
m576-1	GTVFDSSKAN	GGPATFPLSQ	VIPGWTEGVR	LLKEGGEATF	YIPSNLAYRE	QGAGEKIGPN
	190	200	210	220	230	240
	250	260	270			
g576-1.pep	ATLVFDVKLV	KIGAPENAPA	KQPDQVDIKK	VN		
m576-1	ATLVFDVKLV	KIGAPENAPA	KQPDQVDIKK	VN		
	250	260	270			

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3067>:

a576-1.seq
 1 ATGAACACCA TTTTCAAAAT CAGCGCACTG ACCCTTTCCG CCGCTTTGGC
 51 ACTTTCCGCC TGCGGCAAAA AAGAAGCCGC CCCCGCATCT GCATCCGAAC
 101 CTGCCGCCGC TTCTTCCGCG CAGGGCGACA CCTCTTCGAT CGGCAGCAGC
 151 ATGCAGCAGG CAAGCTATGC GATGGGCGTG GACATCGGAC GCTCCCTGAA
 201 GCAAATGAAG GAACAGGGCG CGGAAATCGA TTTGAAAGTC TTTACCGAAG
 251 CCATGCAGGC AGTGTATGAC GGCAAAGAAA TCAAAATGAC CGAAGAGCAG
 301 GCTCAGGAAG TCATGATGAA ATTCCTTCAG GAACAACAGG CTAAGCCGT
 351 AGAAAAACAC AAGGCGGACG CGAAGGCCAA TAAAGAAAAA GGCGAAGCCT
 401 TTCTGAAAGA AAATGCCGCC AAAGACGGCG TGAAGACCAC TGCTTCCGGC
 451 CTGCAATACA AAATCACCAA ACAGGGCGAA GGCAAACAGC CGACCAAGA
 501 CGACATCGTT ACCGTGGAAT ACGAAGGCCG CCTGATTGAC GGTACGGTAT
 551 TCGACAGCAG CAAAGCCAAC GCGGCGCGG TCACCTTCCC TTTAGCCAA
 601 GTGATTCTGG GTTGACCGA AGGCGTACAG CTTCTGAAAG AAGGCGCGA
 651 AGCCACGTTC TACATCCCGT CCAACCTTGC CTACCGCGAA CAGGGTGGCG

701 GCGACAAAAT CGGCCCGAAC GCCACTTTGG TATTTGATGT GAAACTGGTC
 751 AAAATCGGCG CACCCGAAAA CGCGCCCGCC AAGCAGCCGG CTCAAGTCGA
 801 CATCAAAAAA GTAAATTAA

This corresponds to the amino acid sequence <SEQ ID 3068; ORF 576-1.a>:

a576-1.pep
 1 MNTIFKISAL TLSAALALSA CGKKEAAPAS ASEPAASSA QGDTSSIGST
 51 MQQASYAMGV DIGRSLKQMK EQGAEIDLKV FTEAMQAVYD GKEIKMTEEQ
 101 AQEVMKFLQ EQQAKAVEKH KADAKANKEK GEAFLENAA KDGVKTTASG
 151 LQYKITKQGE GKQPTKDDIV TVEYEGRLID GTVFDSSKAN GGPVTFPLSQ
 201 VILGWTEGVQ LLKEGGEATF YIPSNLAYRE QGAGDKIGPN ATLVFDVKLV
 251 KIGAPENAPA KQPAQVDIKK VN*

a576-1/m576-1 ORFs 576-1 and 576-1.a showed a 99.6% identity in 272 aa overlap

	10	20	30	40	50	60
a576-1.pep	MNTIFKISAL	TLSAALALSA	CGKKEAAPAS	ASEPAASSA	QGDTSSIGST	MQQASYAMGV
m576-1	MNTIFKISAL	TLSAALALSA	CGKKEAAPAS	ASEPAASSA	QGDTSSIGST	MQQASYAMGV
	10	20	30	40	50	60
	70	80	90	100	110	120
a576-1.pep	DIGRSLKQMK	EQGAEIDLKV	FTEAMQAVYD	GKEIKMTEEQ	AQEVMKFLQ	EQQAKAVEKH
m576-1	DIGRSLKQMK	EQGAEIDLKV	FTEAMQAVYD	GKEIKMTEEQ	AQEVMKFLQ	EQQAKAVEKH
	70	80	90	100	110	120
	130	140	150	160	170	180
a576-1.pep	KADAKANKEK	GEAFLENAA	KDGVKTTASG	LQYKITKQGE	GKQPTKDDIV	TVEYEGRLID
m576-1	KADAKANKEK	GEAFLENAA	KDGVKTTASG	LQYKITKQGE	GKQPTKDDIV	TVEYEGRLID
	130	140	150	160	170	180
	190	200	210	220	230	240
a576-1.pep	GTVFDSSKAN	GGPVTFPLSQ	VILGWTEGVQ	LLKEGGEATF	YIPSNLAYRE	QGAGDKIGPN
m576-1	GTVFDSSKAN	GGPVTFPLSQ	VILGWTEGVQ	LLKEGGEATF	YIPSNLAYRE	QGAGDKIGPN
	190	200	210	220	230	240
	250	260	270			
a576-1.pep	ATLVFDVKLV	KIGAPENAPA	KQPAQVDIKK	VNX		
m576-1	ATLVFDVKLV	KIGAPENAPA	KQPAQVDIKK	VNX		
	250	260	270			

919 gnm43.seq

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3069>:

m919.seq
 1 ATGAAAAAAT ACCTATTCCG CGCCGCCCTG TACGGCATCG CCGCCGCCAT
 51 CCTCGCCGCC TGCCAAAGCA AGAGCATCCA AACCTTTCCG CAACCCGACA
 101 CATCCGTCAT CAACGGCCCG GACCGGCCGG TCGGCATCCC CGACCCCGCC
 151 GGAACGACGG TCGCGGCGCG CGGGGCCGTC TATACCGTTG TACCGCACCT
 201 GTCCCTGCCC CACTGGGCGG CGCAGGATT CGCCAAAAGC CTGCAATCCT
 251 TCCGCCTCGG CTGCGCCAAT TTGAAAACC GCCAAGGCTG GCAGGATGTG
 301 TGCGCCCAAG CCTTTCAAAC CCGGTCCAT TCCTTTCAGG CAAAACAGTT
 351 TTTTGAACGC TATTTCACGC CGTGGCAGGT TGCAAGCAAC GGAAGCCTTG
 401 CCGGTACGGT TACCGCTAT TACGAACCGG TGCTGAAGGG CGACGACAGG

```

451 CGGACGGCAC AAGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCTGCGG GTTGCGGAG CGGAAAAGCC CTGTCCGCA
551 TCAGGCAGAC GGGAAAAAAC AGCGGCACAA TCGACAATAC CGGCGGCACA
601 CATACCGCCG ACCTCTCCCG ATTCCCACAT ACCGCGCGCA CAACAGCAAT
651 CAAAGGCAGG TTTGAAGGAA GCCGCTTCTT CCCTACCAC ACGCGCAACC
701 AAATCAACGG CGGCGCGCTT GACGGCAAAG CCCCGATACT CGGTTACGCC
751 GAAGACCTG TCGAACTTT TTTATGCAC ATCCAAGGCT CGGGCCGTCT
801 GAAAACCCCG TCCGGCAAAT ACATCCGCAT CGGCTATGCC GACAAAAACG
851 AACATCCyTA CGTTTCCATC GGACGTATA TGGCGGATAA GGGCTACCTC
901 AAATCGGAC AAACCTCCAT GCAGGGCATT AAGTCTTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC AGCAATGACG GCCCTGTCCG CGCACTGGGC
1051 ACGCCGCTGA TGGGGGAATA TGCCGGCGCA GTCGACCGGC ACTACATTAC
1101 CTTGGGTGCG CCTTATTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCGTATTATG GCGCAGGATA CCGGCAGCGC GATTAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGATAC GCGGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGATATGT CTGGCAGCTC CTACCCAACG
1301 GTATGAAGCC CGAATACCGC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3070; ORF 919>:

m919.pep

```

1  MKKYLFRAL YGIAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51  GTTVGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGGDR
151 RTAQAQFPIY GIPDDFISVP LPAGLRSGKA LVRIQTGKN SGTIDNTGGT
201 HTADLSRFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KSYMQRNPOR LAEVLGQNPS YIFFRELAYS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 3071>:

g919.seq

```

1  ATGAAAAAAC ACCTGCTCCG CTCCGCCCTG TACGGcatCG CCGCCgcat
51  CctcgCCGCC TGCCAAAgca gGAGCATCCA AACCTTTCCG CAACCCGACA
101 CATCCGTCAT CAACGGCCCCG GACCGGCCGG CCGGCATCCC CGACCCCGCC
151 GGAACGACGG TTGCCGCGCG CGGGGCCGTC TATACCGTTG TGCCGCACCT
201 GTCCATGCCC CACTGGGCGG CGCaggATTT TGCCAAAAGC CTGCAATCCT
251 TCCGCCTCGG CTGCGCCAAT TTGAAAACC GCCAAGGCTG GCAGGATGTG
301 TGCGCCCAAG CCTTTCAAAC CCCCGTGCAT TCCTTTCAG CAAAGcGgTT
351 TTTTGAACGC TATTTACGC cgtGGCagg tgcaggcaAC GGAAGcCTTG
401 CaggtaaggT TACCGGCTAT TACGAACCGG TGCTGAAGGG CGACGGCAGG
451 CGGACGGAAC GGGCCCGCTT CCCGATTAC GGTATTCCCG ACGATTTTAT
501 CTCCGTCCCC CTGCTGCGG GTTGCGGGG CGGAAAAAAC CTTGTCCGCA
551 TCAGGCAGac ggGGAANAAC AGCGGCACGA TCGACAATGC CGGCGGCACG
601 CATACGCCCC ACCTCTCCCC ATTCCCACAT ACCGCGCGCA CAACGGcaat
651 caaaGGCAGG TTTGAaggAA GCCGCTTCTT CCCTTACCAC ACGCGCAACC
701 AAAtcaacGG CGGCGcgctT GACGGCAAag cccCATCCT CggttacgcC
751 GAagaccCcG tcgaacttTT TTTATGCAC AtccaaggCT CGGGCCGCCT
801 GAAAACCCcg tccggcaaat acatCCGCAt cggaTatgcc gacAAAAACG
851 AACAtccgTa tgtttccatc ggACGctaTA TGGCGGACAA AGGCTACCTC
901 AAGctcgggc agACCTCGAT GCAGGgcatc aaagcCTATA TGCGGCAAAA
951 TCCGCAACGC CTCGCCGAAG TTTTGGGTCA AAACCCAGC TATATCTTTT
1001 TCCGCGAGCT TGCCGGAAGC GGCAATGAGG GCCCCGTCCG CGCACTGGGC
1051 ACGCCACTGA TGGGGGAATA CGCCGGCGCA ATCGACCGGC ACTACATTAC
1101 CTTGGGCGCG CCTTATTG TCGCCACCGC CCATCCGGTT ACCCGCAAAG
1151 CCCTCAACCG CCGTATTATG GCGCAGGATA CAGGCAGCGC GATCAAAGGC
1201 GCGGTGCGCG TGGATTATTT TTGGGGTTAC GCGGACGAAG CCGGCGAACT
1251 TGCCGGCAAA CAGAAAACCA CGGATACGT CTGGCAGCTC CTGCCCCAACG
1301 GCATGAAGCC CGAATACCGC CCGTGA

```

ORF 919 shows 95.9 % identity over a 441 aa overlap with a predicted ORF (ORF 919.ng) from *N. gonorrhoeae*:

[illegible]

```

              430      440
m919.pep      QKTTGYVWQLLPNGMKPEYRPX
              |||||
g919           QKTTGYVWQLLPNGMKPEYRPX
              430      440

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 3073>:

```

a919.seq
1  ATGAAAAAAT  ACCTATTCGG  CGCCGCCCTG  TGC GG CATCG  CCGCCGCCAT
51  CCTCGCCGCC  TGCCAAAGCA  AGAGCATCCA  AACCTTTCCG  CAACCCGACA
101 CATCCGTCAT  CAACGGCCCG  GACCGGCCGG  TCGGCATCCC  CGACCCCGCC
151 GGAACGACGG  TCGGCGCGGG  .CGGGGCGGTT  TATACCGTTG  TGCCGCACCT
201 GTCCCTGCCC  CACTGGGCGG  CGCAGGATTT  CGCCAAAAGC  CTGCAATCCT
251 TCCGCCTCGG  CTGCGCCAAT  TTGAAAACCC  GCCAAGGCTG  GCAGGATGTG
301 TGC GCCCAAG  CCTTTCAAAC  CCCCGTCCAT  TCCGTTCAAG  CAAAACAGTT
351 TTTTGAACGC  TATTTACGCG  CGTGGCAGGT  TGCAGGCAAC  GGAAGCCTTG
401 CCGGTACGGT  TACCGGCTAT  TACGAGCCGG  TGCTGAAGGG  CGACGACAGG
451 CGGACGCGAC  AAGCCCGCTT  CCCGATTAC  GGTATTTCCG  ACGATTTTAT
501 CTCCGTCCCC  CTGCCTGCCG  GTTTGCGGAG  CGGAAAAGCC  CTTGTCCGCA
551 TCAGGCAGAC  GGGAAAAAAC  AGCGGCACAA  TCGACAATAC  CGGCGGCACA
601 CATACCGCCG  ACCTCTCCCA  ATTCCCCATC  ACTGCGCGCA  CAACGCACAT
651 CAAAGGCAGG  TTTGAAGGAA  GCCGCTTCCT  CCCCTACCAC  ACGCGCAACC
701 AAATCAACGG  CGGCGCGCTT  GACGGCAAAG  CCCCATACT  CGGTTACGCC
751 GAAGACCCCG  TCGAACTTTT  TTTTATGCAC  ATCCAAGGCT  CGGGCCGTCT
801 GAAAACCCCG  TCCGGCAAAT  ACATCCGCAT  CGGCTATGCC  GACAAAACG
851 AACATCCCTA  CGTTTCCATC  GGACGCTATA  TGGCGGACAA  AGGCTACCTC
901 AAGCTCGGGC  AGACCTCGAT  GCAGGCGATC  AAAGCTATA  TGCAGCAAAA
951 CCCGCAACGC  CTCGCCAAG  TTTTGGGGCA  AAACCCAGC  TATATCTTTT
1001 TCCGAGAGCT  TACCGGAAGC  AGCAATGACG  GCCCTGTCGG  CGCACTGGGC
1051 ACGCCGCTGA  TGGGCGAGTA  CGCCGGCGCA  GTCGACCGGC  ACTACATTAC
1101 CTTGGGCGCG  CCCTTATTTG  TCGCCACCGC  CCATCCGGTT  ACCCGCAAAG
1151 CCCTCAACCG  CCTGATTATG  GCGCAGGATA  CCGGACGCGC  GATTAAAGGC
1201 GCGGTGCGCG  TGGATTATTT  TTGGGATAC  GGCGACGAAG  CCGCGCAAT
1251 TGCCGGCAAA  CAGAAAACCA  CGGGATATGT  CTGGCAGCTT  CTGCCCAACG
1301 GTATGAAGCC  CGAATACCGC  CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 3074; ORF 919.a>:

```

a919.pep
1  MKKYLFRAL  CGIAAAILAA  CQSKSIQTFP  QPDTSVINGP  DRPVGIPDPA
51  GTTVGGGGAV  YTVVPHLSLP  HWAQAQDFAKS  LQSFRLGCAN  LKNRQGWQDV
101 CAQAFQTPVH  SVQAKQFFER  YFTPWQVAGN  GSLAGTVTGY  YEPVLKGDDR
151 RTAQAQRFPIY  GIPDDFISVP  LPAGLRSGKA  LVRIRQTGKN  SGTIDNTGGT
201 HTADLSQFPI  TARTTAIKGR  FEGRFLPYH  TRNQINGGAL  DGKAPILGYA
251 EDPVELFFMH  IQSGRLKTP  SGKYIRIGYA  DKNEHPYVSI  GRYMADKGYL
301 KLGQTSMQGI  KAYMQQNPQR  LAEVLGQNPS  YIFFRELTGS  SNDGPVGALG
351 TPLMGEYAGA  VDRHYITLGA  PLFVATAHPV  TRKALNRLIM  AQDTGSAIKG
401 AVRVDYFWGY  GDEAGELAGK  QKTTGYVWQL  LPNGMKPEYR  P*

```

m919/a919 ORFs 919 and 919.a showed a 98.6% identity in 441 aa overlap

```

              10      20      30      40      50      60
m919.pep      MKKYLFRALYGI AAAA ILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
              |||||
a919           MKKYLFRALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
              10      20      30      40      50      60

              70      80      90      100     110     120
m919.pep      YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
              |||||
a919           YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
              70      80      90      100     110     120

```

	130	140	150	160	170	180
m919.pep	YFTPWQVAGNGSLAGTVTGYIEPVLKGD	DRRTAQA	RFPIYGIPDDFISVPLPAGLRSGKA			
a919	YFTPWQVAGNGSLAGTVTGYIEPVLKGD	DRRTAQA	RFPIYGIPDDFISVPLPAGLRSGKA			
	130	140	150	160	170	180
	190	200	210	220	230	240
m919.pep	LVRIRQTGKNSGTIDNTGGTHTADLSRFP	ITARTTAIKGRFEGSRFLPYHTRNQINGGAL				
a919	LVRIRQTGKNSGTIDNTGGTHTADLSQFP	ITARTTAIKGRFEGSRFLPYHTRNQINGGAL				
	190	200	210	220	230	240
	250	260	270	280	290	300
m919.pep	DGKAPILGYAEDPVELFFMH	IQSGRLKTPSGKYIRIGYADKNEHPYV	SIGRYMADKGYL			
a919	DGKAPILGYAEDPVELFFMH	IQSGRLKTPSGKYIRIGYADKNEHPYV	SIGRYMADKGYL			
	250	260	270	280	290	300
	310	320	330	340	350	360
m919.pep	KLGGTSMQGIKSYMQRNPQRLAEVLGQN	PSYIFFRELAGSSNDGPVGALGTPLMGEYAGA				
a919	KLGGTSMQGIKAYMQQN	PQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA				
	310	320	330	340	350	360
	370	380	390	400	410	420
m919.pep	VDRHYITLGAPLFVATAHPVTRKALNRL	IMAQDTGSAIKGAVRVDFWGYGDEAGELACK				
a919	VDRHYITLGAPLFVATAHPVTRKALNRL	IMAQDTGSAIKGAVRVDFWGYGDEAGELACK				
	370	380	390	400	410	420
	430	440				
m919.pep	QKTTGYVWQLLPNGMKPEYRPX					
a919	QKTTGYVWQLLPNGMKPEYRPX					
	430	440				

121 and 121-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3075>:

m121.seq

```

1  ATGGAACAC  AGCTTACAT  CGGCATCATG  TCGGGAACCA  GCATGGACGG
51  GCGGATGCC  GACTGATAC  GGATGGACGG  CGGCAAATGG  CTGGGCGCGG
101 AAGGGCACGC  CTTTACCCCC  TACCCCGGCA  GGTTACGCCG  CCAATTGCTG
151 GATTTCAGG  ACACAGGCGC  AGACGAACTG  CACCGCAGCA  GGATTTGTG
201 GCAAGAACTC  AGCCGCCTAT  ATGCGCAAAC  CGCCGCCGAA  CTGCTGTGCA
251 GTCAAAACCT  CGCACCGTCC  GACATTACCG  CCCTCGGCTG  CCACGGGCAA
301 ACCGTCCGAC  ACGCGCCGGA  ACACGGTTAC  AGCATACAGC  TTGCCGATTT
351 GCCGTGCTG  GCGxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx
401 xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx
451 xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx
501 xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx
551 xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx  xxxxxxxxxx
601 xxxxxxGAG  TTCCTTACGA  CAAAACGGT  GCAAAGTCGG  CACAAGGCAA
651 CATATTGCCG  CAACTGCTCG  ACAGGCTGCT  CGCCACCCG  TATTTCGCAC
701 AACGCCACCC  TAAAGCAGC  GGGCGGAAC  TGTTTGCCAT  AAATTGGCTC
751 GAAACCTACC  TTGACGGCGG  CGAAAACCGA  TACGACGTAT  TGCGGACGCT
801 TTCCCGTTT  ACCGCGCAA  CCGTTTGCGA  CGCCGTCTCA  CACGCAGCGG
851 CAGATGCCCG  TCAATGTAC  ATTGCGCAG  GCGGCATCCG  CAATCCTGTT
901 TTAATGGCG  ATTGGCAGA  ATGTTTCGG  ACACGCGTTT  CCTGACACAG
951 CACCGCCGAC  CTGAACCTCG  ATCCGCAATG  GGTGGAAGCC  GCCGnATTG
1001 CGTGGTTGG  GCGGTGTTG  ATTAATCGCA  TTCCCGGTAG  TCCGCACAAA

```

1051 GCAACCGGCG CATCCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG
1101 A

This corresponds to the amino acid sequence <SEQ ID 3076; ORF 121>:

m121.pep

1 METQLYIGIM SGTSMGADA VLIRMDGGKW LGAEGHAFTP YPGRLRRQLL
51 DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL AXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
151 XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX XXXXXXXXXXXX
201 XXQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRTLRSF TAQTVCDAYS HAAADARQMY ICDGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3077>:

g121.seq

1 ATGGAAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51 GCGGGATGCC GTGCTGGTAC GGATGGACGG CGGCAATGCG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCTGACC GGTGCGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCAC AGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCGCGCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATACAGC TTGCCGATTT
351 GCGCGTGTG GCGGAACTGa cgcggatttt TACCGTCgyc gacttcCGCA
401 GCGCGACCT TGCTGCCGCG GGacaAGGTG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CCGCGGGATT GCCAACATCA GCGTACTCCC CCCCGGCGCA CCCGCTTTCG
551 GCTTCGACAC AGGGCCCGGC AATATGCTGA TGGAcgcgtg gacgcaggca
601 cacTGGcagc TGCCTTACGA CAAAaAcggt gcAAAGgcyg cacAAGGCAA
651 catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
701 AACCCcacc aaAAAGCAGC GGgcGCGaac TgtttgcccT AAattggctc
751 gaaacctAcc ttgacggcgg cgaaaaccga tacgacgtat tgcggacgct
801 ttcccgatc accgcgcaaa ccgTttggga cgccgtctca CACGACGCG
851 CAGATGCCCC TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGCGCG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATttg
1001 cgtggttggC GCGGTGTTGG ATTAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

This corresponds to the amino acid sequence <SEQ ID 3078; ORF 121.ng>:

g121.pep

1 METQLYIGIM SGTSMGADA VLVRMDGGKW LGAEGHAFTP YPDRLRRLKLL
51 DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPC DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTGPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLRSF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae*:

m121/g121

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGADAVLIRMDGGKWLGAEHGAFTPYPGRLRRQLLDLQDTGADEL					
	: : : : :					
g121	METQLYIGIMSGTSMGADAVLVRMDGGKWLGAEHGAFTPYPDRLRRLKLLDQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEGYSIQLADLPLL					
	: : : : :					

```

g121      HRSRMLSQELSRLYAQTAAELLCSONLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
              70      80      90      100      110      120
              130      140      150      160      170      180
m121.pep  AXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
              :
g121      AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
              130      140      150      160      170      180
              190      200      210      220      230      240
m121.pep  XXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST
              :
g121      PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST
              190      200      210      220      230      240
              250      260      270      280      290      300
m121.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDASHAAADARQMYICDGGIRNPV
              :
g121      GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDASHAAADARQMYICGGIRNPV
              250      260      270      280      290      300
              310      320      330      340      350      360
m121.pep  LMADLAECFGTRVSLHSTADNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
              :
g121      LMADLAECFGTRVSLHSTAEINLDPQWVEAAAFWLAACWINRIPGSPHKATGASKPCIL
              310      320      330      340      350      360

m121.pep  XAGYXX
              |||||
g121      GAGYXX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3079>:

```

a121.seq
1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAACGT CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCGCTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTACAGC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGCGCCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGCGGGGATT GCCAACATCA GCGTACTCCC CCCCAGCGCA CCCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTCGCAC
701 AACCCACCC TAAAGCACG GGGCGCGAAC TGTTCGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TCGGACGCT
801 TTCCCGATTC ACCGCGCAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGCGC CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3080; ORF 121.a>:

```

a121.pep
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPGP NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV

```

301 LMADLAECFG TRVSLHSTAE LNLDPOWVEA AAFAWMAACW VNRIPGSPHK
 351 ATGASKPCIL GAGYYY*

m121/a121 ORFs 121 and 121.a showed a 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSM	DGADAVLIRMDGGK	WLGAEHGAFTYPY	PGRLRRQLLDLQD	TGADEL	
a121	METQLYIGIMSGTSM	DGADAVLIRMDGGK	WLGAEHGAFTYPY	PGRLRRKLLDLQD	TGADEL	
	70	80	90	100	110	120
m121.pep	HRSRILSQELSR	LYAQTAELLCSQ	NLAPSDITALGCH	GQTVRHAPEHGY	SIQLADLP	LL
a121	HRSRMLSQELSR	LYAQTAELLCSQ	NLAPSDITALGCH	GQTVRHAPEHSY	SVQLADLP	LL
	130	140	150	160	170	180
m121.pep	AXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX
a121	AERTQIFTVGDF	RSRDLAAGGQGA	PLVPAFHEALFR	DDRETRAVLNIG	GGIANISVLP	PPDA
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX	XXXXXXXXXXXXXX
a121	PAFGFDTGPGN	MLMDAWMQAHW	QLPYDKNGAKAA	QGNILPQLLDRL	LAHPYFAQPH	PKST
	250	260	270	280	290	300
m121.pep	GRELFAINWLET	YLDGGENRYDVL	RTLSRFTAQTV	CDVASHAAADAR	QMYICDGGIR	NPV
a121	GRELFAINWLET	YLDGGENRYDVL	RTLSRFTAQTV	FDVASHAAADAR	QMYICGGGIR	NPV
	310	320	330	340	350	360
m121.pep	LMADLAECFGTR	VSLHSTADLNLD	POWVEAAXFAW	LAACWINRIPG	SPHKATGASK	KPCIL
a121	LMADLAECFGTR	VSLHSTAE LNLD	POWVEAAAFAW	MAACWVNRIP	GSPHKATGASK	KPCIL
m121.pep	XAGYYYYX					
a121	GAGYYYYX					

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3081>:

m121-1.seq

1	ATGGAACAC	AGCTTTACAT	CGGCATCATG	TCGGGAACCA	GCATGGACGG
51	GGCGGATGCC	GTA CTGATAC	GGATGGACGG	CGGCAAATGG	CTGGGCGCGG
101	AAGGGCACGC	CTTTACCCCC	TACCCCGGCA	GGTTACGCCG	CCAATTGCTG
151	GATTTGCAGG	ACACAGGCGC	AGACGAACTG	CACCGCAGCA	GGATTTTGTC
201	GCAAGAACTC	AGCCGCCTAT	ATGCGCAAAC	CGCCGCCGAA	CTGCTGTGCA
251	GTCAAAACCT	CGCACCCTCC	GACATTACCG	CCCTCGGCTG	CCACGGGCAA
301	ACCGTCCGAC	ACGCGCCGGA	ACACGGTTAC	AGCATACAGC	TTGCCGATTT
351	GCCGCTGCTG	GCGGAACGGA	CGCGGATTTT	TACCGTCGGC	GACTTCCGCA
401	GCCGCGACCT	TGCGGCCGGC	GGACAAGGCG	CGCCACTCGT	CCCCGCCTTT
451	CACGAAGCCC	TGTTCCGCGA	CAACAGGGAA	ACACGCGCGG	TACTGAACAT
501	CGCGGGGATT	GCCAAACATCA	GCGTACTCCC	CCCCGACGCA	CCCCCCTTCG
551	GCTTCGACAC	AGGGCCGGGC	AATATGCTGA	TGGACGCGTG	GACGCAGGCA
601	CACTGGCAGC	TTCTTACGA	CAAAAACGGT	GCAAAGGCGG	CACAAGGCAA
651	CATATTGCCG	CAACTGCTCG	ACAGGCTGCT	CGCCACCCCG	TATTTGCGAC
701	AACCCACCCC	TAAAAGCACG	GGGCGCGAAC	TGTTTGCCCT	AAATTGGCTC
751	GAAACCTACC	TTGACGGCGG	CGAAAACCGA	TACGACGTAT	TGCGGACGCT
801	TTCCCGTTTT	ACCGCGCAAA	CCGTTGCGA	CGCGTCTCA	CACGCAGCGG


```

851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTTG
1001 CGTGGTTGGC GCGGTGTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG ANCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3082; ORF 121-1>:

```

m121-1.pep
1  METQLYIGIM SGTSMGADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRQLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPHEGY SIQLADLPLL AERTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTPGP NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVCDAYS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

m121-1/g121 ORFs 121-1 and 121.ng showed a 95.6% identity in 366 aa overlap

	10	20	30	40	50	60
m121-1.pep	METQLYIGIMSGTSMGADAVLIRMDGGKW	LGAEGHAFTPYPGRLLRRQLL	DLQDTGADEL			
g121	METQLYIGIMSGTSMGADAVLVRMDGGKW	LGAEGHAFTPYPDRLLRRKLL	DLQDTGTDEL			
	10	20	30	40	50	60
	70	80	90	100	110	120
m121-1.pep	HRSRILSQELSRLYAQTAAEL	LLCSQNLAPSDITALGCHGQ	TVRHAPHEGYSIQLADLPLL			
g121	HRSRILSQELSRLYAQTAAEL	LLCSQNLAPCDITALGCHGQ	TVRHAPHEGYSIQLADLPLL			
	70	80	90	100	110	120
	130	140	150	160	170	180
m121-1.pep	AERTRIFTVGDFRSRDLAAGG	QGAPLVPAFHEALFRDNRE	TRAVLNIGGIANISVLPPDA			
g121	AELTRIFTVGDFRSRDLAAGG	QGAPLVPAFHEALFRDDRE	TRVVLNIGGIANISVLPPGA			
	130	140	150	160	170	180
	190	200	210	220	230	240
m121-1.pep	PAFGFDTPGPNMLMDAWTQA	HWQLPYDKNGAKAAQGNILP	QLLDRLLAHPYFAQPHPKST			
g121	PAFGFDTPGPNMLMDAWTQA	HWQLPYDKNGAKAAQGNILP	QLLGRLLAHPYFSQPHPKST			
	190	200	210	220	230	240
	250	260	270	280	290	300
m121-1.pep	GRELFALNWLETYLDGGENR	YDVLRTLSRFTAQTVCDAYS	HAAADARQMYICGGGIRNPV			
g121	GRELFALNWLETYLDGGENR	YDVLRTLSRFTAQTVWDAYS	HAAADARQMYICGGGIRNPV			
	250	260	270	280	290	300
	310	320	330	340	350	360
m121-1.pep	LMADLAECFGTRVSLHSTAD	LNLDPQWVEAAXFAWLAACW	INRIPGSPHKATGASKPCIL			
g121	LMADLAECFGTRVSLHSTAE	LNLDPQWVEAAAFWLAACW	INRIPGSPHKATGASKPCIL			
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
g121	GAGYYYYX					

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3083>:

```

a121-1.seq
1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG

```

```

51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
151 GATTTGCAGG ACACAGGCGC GGACGAACTG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCCTGT ACGCGCAAA CCGCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGGGCCGGC GGACAAGGCG CGCCGCTCGT CCCCCTTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCACGCA CCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT CCAAAGCGCG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTGCGAC
701 AACCCACCCC TAAAGCACG GGGCGCGAAC TGTTTGCCTT AAATTGGCTC
751 GAAACCTACC TTGACGCGCG CGAAAACCGA TACGACGTAT TCGGGACGCT
801 TTCCCGATT CCGCGCAAA CCGTTTTCGA CGCCGCTCTA CACGACGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCAAACC GTGTATTCTG GCGCGGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 3084; ORF 121-1.a>:

a121-1.pep

```

1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCQNLAAPS DITALGCHGQ
101 TVRHAPHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAFLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMDLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYY*

```

m121-1/a121-1 ORFs 121-1 and 121-1.a showed a 96.4% identity in 366 aa overlap

	10	20	30	40	50	60
m121-1.pep	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEHGAFTPYPGRLRRQLDLQDTGADEL					
a121-1	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEHGAFTPYPGRLRRQLDLQDTGADEL					
	10	20	30	40	50	60
m121-1.pep	HRSRILSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAPHEHYSIQLADLPLL					
a121-1	HRSRILSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAPHEHYSIQLADLPLL					
	70	80	90	100	110	120
m121-1.pep	HRSRILSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAPHEHYSIQLADLPLL					
a121-1	HRSRILSQELSRLYAQTAAELLCSQNLAAPSDITALGCHGQTVRHAPHEHYSIQLADLPLL					
	70	80	90	100	110	120
m121-1.pep	AERTQIFTVGDFRSRDLAAGGQGAFLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA					
a121-1	AERTQIFTVGDFRSRDLAAGGQGAFLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
m121-1.pep	AERTQIFTVGDFRSRDLAAGGQGAFLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA					
a121-1	AERTQIFTVGDFRSRDLAAGGQGAFLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
m121-1.pep	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
a121-1	PAFGFDTPGPNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
m121-1.pep	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
a121-1	PAFGFDTPGPNMLMDAWMQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
m121-1.pep	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICGGGIRNPV					
a121-1	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFCDVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
m121-1.pep	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDVSHAAADARQMYICGGGIRNPV					
a121-1	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFCDVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300

```

          310      320      330      340      350      360
m121-1.pep  LMADLAECFGTRVSLHSTADLNLDLPQWVEAAAFAWLAACWINRIPGSPHKATGASKPCIL
          |||||:|||||:|||||:|||||:|||||:|||||
a121        LMADLAECFGTRVSLHSTAE LNLDLPQWVEAAAFAWMAACWVNRI PGSPHKATGASKPCIL
          310      320      330      340      350      360

m121-1.pep  XAGYYYY
          |||||
a121        GAGYYYY

```

128 and 128-1

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3085>:

```

m128.seq (partial)
  1  ATGACTGACA  ACGCACTGCT  CCATTTGGGC  GAAGAACCCC  GTTTTGATCA
 51  AATCAAAACC  GAAGACATCA  AACCCGCCCT  GCAAACCGCC  ATCGCCGAAG
101  CGCGCGAACA  AATCGCCGCC  ATCAAAGCCC  AAACGCACAC  CGGCTGGGCA
151  AACACTGTCG  AACCCCTGAC  CGGCATCACC  GAACGCGTCG  GCAGGATTTG
201  GGGCGTGGTG  TCGCACCTCA  ACTGCGTCGC  CGACACGCCC  GAACTGCGCG
251  CCGTCTATAA  CGAACTGATG  CCGAAATCA  CCGTCTTCTT  CACCGAAATC
301  GGACAAGACA  TCGAGCTGTA  CAACCGCTTC  AAAACCATCA  AAAATTCCCC
351  CGAATTCGAC  ACCCTCTCCC  CCGCACAAAA  AACCAAACTC  AACCAC
   1  TACGCCAGCG  AAAAACTGCG  CGAAGCCAAA  TACGCGTTCA  GCGAAACCGA
 51  wGTCAAAAA  TAyTTCCCyG  TCGGCAAwGT  ATTAACCGGA  CTGTTCCGCC
101  AAmTCAAAAA  ACTmTACGGC  ATCGGATTTA  CCGAAAAAAC  yGTCCCGTC
151  TGGCACAAAG  ACGTGCGCTA  TTKTGAATTG  CAACAAAACG  GCGAaMCCAT
201  AGGCGGCGTT  TATATGGATT  TGTACGCACG  CGAAGGCAAA  CGCGGCGGCG
251  CGTGGATGAA  CGACTACAAA  GGCCGCCGCC  GTTTTTCAGA  CGGCACGCTG
301  CAAYTGCCCA  CCGCTACCT  CGTCTGCAAC  TTCGCCCCAC  CCGTCGGCGG
351  CAGGGAAGCC  CGCyTAGGCC  ACGACGAAAT  CCTCATCTCT  TTCCACGAAA
401  CCGGACACGG  GCTGCACCC  CTGCTTACCC  AAGTGGACGA  ACTGGGCGTA
451  TCCGGCATCA  ACGGCGTAKA  ATGGGACGCG  GTCGAAGTGC  CCAGCCAGTT
501  TATGGAATA  TTCGTTTGGG  AATACAATGT  CTTGGCACAA  mTGTCAGCCC
551  ACGAAGAAAC  CGGcgTTCCC  yTGCCGAAAG  AACTCTTsGA  CAAAwTGCTC
601  GCCGCCAAAA  ACTTCCAAaG  CGGCATGTTC  yTsGTCCGGC  AAwTGAGTT
651  CGCCCTCTTT  GATATGATGA  TTTACAGCGA  AGACGACGAA  GGCCGTCTGA
701  AAAACTGGCA  ACAGGTTTTA  GACAGCGTGC  GCAAAAAAGT  CGCCGTCTATC
751  CAGCCGCCCG  AATACAACCG  CTTCCGCTTG  AGCTTCGGCC  ACATCTTCGC
801  AGGCGGCTAT  TCCGCAGCTn  ATTACAGCTA  CGCGTGGGCG  GAAGTATTGA
851  GCGCGGACGC  ATACGCCGCC  TTTGAAGAAA  GCGACGATGT  CGCCGCCACA
901  GGCAAACGCT  TTTGGCAGGA  AATCCTCGCC  GTCGGGGnAT  CGCGCAGCGG
951  nGCAGAATCC  TTCAAAGCCT  TCCGCGGCCG  CGAACCGAGC  ATAGACGCAC
1001 TCTTGCGCCA  CAGCGGTTTC  GACAACGCGG  TCTGA

```

This corresponds to the amino acid sequence <SEQ ID 3086; ORF 128>:

```

m128.pep (partial)
  1  MTDNALLHLG  EEPFRDQIKT  EDIKPALQTA  IAEAREQIAA  IKAQTHTGWA
 51  NTVEPLTGIT  ERVGRIWGVV  SHLNCVADTP  ELRAVYNELM  PEITVFFTEI
101  GDIELYNRF  KTIKNSPEFD  TLSPAQKTKL  NH

//

  1  YASEKLREAK  YAFSETXVKK  YFPVGXVLNG  LFAQXKKLYG  IGFTEKTPVP
 51  WHKDVRYXEL  QONGEXIGGV  YMDLYAREGK  RGGAWMNDYK  GRRRFS DGTL
101  QLP TAYLVCN  FAPPVGGREA  RLSHDEILIL  FHETGHGLHH  LLTQVDELGV
151  SGINGVXWDA  VELPSQFMEN  FVWEYNVLAQ  XSAHEETGVP  LPKELXDKXL
201  AAKNFQXGMF  XVRQXEFALF  DMMIYSEDDE  GRLKNWQQVL  DSVRKKVAVI
251  QPPEYNRFAL  SFGHIFAGGY  SAAXYSYAWA  EVLSADAYAA  FEESDDVAAT
301  GKRFWQEILA  VGXSRSGAES  FKAFRGREPS  IDALLRHS GF  DNAV*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3087>:

```

g128.seq
1  atgattgaca acgCActgct ccacttgggc gaagaaccCC GTTTaatca
51  aatccaaacc gaagACatca AACCCGCCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGCA ACGCTTTCCC CCGCACAAAA AACCAAGCTC GATCAGGACC
401 TGCGCGATTT CGTATTGAGC GCGCGGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAA CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA .CGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTATACC AATACGCCGG CAACGCGGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGccaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAACTTCCT GCACGACCTC
901 GCCGCGCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCCGCCCG GAACACCTCG GTCTCGCCGA CCCGAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCAAAACC
1201 ATCGCGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTaca AAGGCCGCGC CCGCTTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAGAAG CGGTTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AacCGGCCAC GGAATGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCGGGCAT CAacggcgtA GAATGGGACG CGGTGGAAT GCCCAGCCAG
1501 TTTATGGAAT ACTTCGTTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551 CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTT GACAAAATGC
1601 TcgcCGCAA AAACCTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTACAGT GAAAGCGACG AATGCCGTCT
1701 GAAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGCGAG CTATTACAGC TACGCATGGG CCGAAGTCCT
1851 cAGACCGGAT GCCTACGCCG CCTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951 gcgCGGGAAT CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTGCTGCGC CAaagcggT TCGACAACGC gGcttgA

```

This corresponds to the amino acid sequence <SEQ ID 3088; ORF 128.ng>:

```

g128.pep
1  MIDNALLHLG EEPFRNQIQT EDIKPAVQTA IAEARGQIAA VKAQHTGWAA
51  NTVRLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 QDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDSLAYAGEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
401 IGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVEPLSQ
501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS ESDECRKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS

```

651 AAESFKAFRG REPSIDALLR QSGFDNAA*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng)
from *N. gonorrhoeae*:

m128/g128

```

..      10      20      30      40      50      60
g128.pep MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQTHGTGWANTVERLTGIT
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          10      20      30      40      50      60
          MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIAKAQTHGTGWANTVEPLTGIT

      70      80      90     100     110     120
g128.pep ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          70      80      90     100     110     120
          ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD

      130     140     150     160     170     180
g128.pep TLSPAQKTKLDHDLRDFVLSGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          130
          //

                                340     350     360
g128.pep                                YAGEKLREAKYAFSETEVKKYFPVGKVLG
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
                                YASEKLREAKYAFSETXVKKYFPVGXVLNG
                                10      20      30

      370     380     390     400     410     420
g128.pep LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          40      50      60      70      80      90
          LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK

      430     440     450     460     470     480
g128.pep GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVDELGV
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          100     110     120     130     140     150
          GRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILTFHETGHGLHLLTQVDELGV

      490     500     510     520     530     540
g128.pep SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQXGMF
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          160     170     180     190     200     210
          SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKMLAAKNFQXGMF

      550     560     570     580     590     600
g128.pep LVRQMEFALFDMMIYSESDCRLKNWQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          220     230     240     250     260     270
          XVRQXEFALFDMMIYSEDEGRKNWQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY

      610     620     630     640     650     660
g128.pep SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRGREPS
m128      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
          280     290     300     310     320     330
          SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSAGAESFKAFRGREPS

```

```

          670      679
g128.pep  IDALLRQSGFDNAAX
          |||||:|||||:
m128      IDALLRHSGFDAVX
          340

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3089>:

```

a128.seq
1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAATC AACACGATC
401 TCGCGGATTT CGTCTCAGC GCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTCCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGCGATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101 CCAATCAAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGGCGCGC TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCGGCGAT CAACGGCGTA GAATGGGACG CAGTCGAAC TCCAGTCAG
1501 TTTATGGAAA ATTTGTTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAATCTTTC GACAAAATGC
1601 TCGCGGCCAA AAATTTCCAA CGCGGAATGT TCCTCGTCCG CCAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCGCGAG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGGAG GAAATCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3090; ORF 128.a>:

```

a128.pep
1  MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTHTGWA
51  NTVFELTGIT ERVGRIWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTYG KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLQP WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTQLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVEFPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

```

551 FALFDMMIYS EDDEGR LKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
 601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRWFQ EILAVGGSRS
 651 AAESFKAFRG REPSIDALLR HSGFDNAA*

m128/a128 ORFs 128 and 128.a showed a 66.0% identity in 677 aa overlap

m128.pep	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT
m128.pep	ERVGRIWGVVSHLNCVADTPELRVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
a128	ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
m128.pep	TLSPAQKTKLNH-----
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128.pep	-----
a128	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV
m128.pep	-----
a128	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMATDPEQVLNLFHDL
m128.pep	-----YASEKLREAKYAFSETXVKKYFPVGX
a128	ARRAKPYAEKD LAEVKAFARESLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK
m128.pep	VLNGLFAQXKKLYGIGFTEKTPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM
a128	VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
m128.pep	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128.pep	ELGVSGINGVXWD AVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAANKFQ
a128	ELGVSGINGVEWD AVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAANKFQ
m128.pep	XGMFXVRQXEFALFDMMIYSEDEGR LKNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF
a128	RGMFLVRQMEFALFDMMIYSEDEGR LKNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF

```

      400      410      420      430      440      450
m128.pep  AGGYSAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRG
          |||||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a128      AGGYSAAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG
          610      620      630      640      650      660

      460      470
m128.pep  REPSIDALLRHSGFDNAVX
          ||||| ||||| ||||| |||||
a128      REPSIDALLRHSGFDNAAX
          670

```

Further work revealed the DNA sequence identified in *N. meningitidis* <SEQ ID 3091>:

```

m128-1.seq
1  ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACCAGATC
401 TGGCGGATTT CGTCTCAGC GCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAACTGGCAA AACTGCAAAAC CGAAGGCGCG CAACTTTCCG CCAAATCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCGGCC
601 GCCCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 AACTACCTC GCCGTATCC AATACGCCGA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAACTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGCAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAAGTTT TAACTTCTT GCACGACCTC
901 GCCCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGCG GAAAGCCTGA ACCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTGCG
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCGC
1151 TCTGGCACA AGACGTGCGC TATTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCG CCGTTTTTCA GACGGCAGCG
1301 TGCAACTGCC CACCGCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCCGCTGAG CCACGACGAA ATCCTCATCC TCTCCACGA
1401 AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CGGTGGAAT GCCCAGCCAG
1501 TTTATGGAAT ATTTGTTTG GGAATACAAT GTCTTGGCAC AAATGTCAGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAATTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAAAA GTCGCCGTCA
1751 TCCAGCCGCC CGAATACAAC CGCTTCGCCT TGAGCTTCGG CCACATCTTC
1801 GCAGCGGCGT ATTCCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCCGCGCG CGGAACCGA GCATAGACGC
2001 ACTCTTGC GC CACAGCGGTT TCGACAACGC GGTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 3092; ORF 128-1>:

```

m128-1.pep.
1  MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTWA
51  NTVEPLTGIT ERVGRIWGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI

```



```

101  GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NHDLRDFVLS GAELPPEQQA
151  ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDAAAPLAGI PEDALAMFAA
201  AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251  KFDNTANIDR TLNALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301  ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351  EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQQNGET
401  IGGVYMDLYA REGKRGGAWM NDYKRRRFS DGTLLQLPTAY LVCNFAPPVG
451  GREARLSHDE ILILFHETGH GLHLLTQVD ELGVSGINGV EWDAVEPLSQ
501  FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551  FALFDMMIYS EDEGRLLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601  AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651  AAESFKAFRG REPSIDALLR HSGFDNAV*

```

The following DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3093>:

g128-1.seq (partial)

```

1  ATGATTGACA ACGCACTGCT CCACTTGGGC GAAGAACCCC GTTTTAATCA
51  AATCAAAACC GAAGACATCA AACCCGCCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTGT
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCG GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGCA ACGCTTTCCC CCGCACAAA AACCAAGCTC GATCACGACC
401 TGCGCGATTT CGTATTGAGC GCGCGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCCG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCGCCAAA CTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGCTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCGC GAACACCTCG GTCTCGCCGA CCCGAGCCG TGGGACTTGA
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTTCTGGCAG GCCTGTTCCG
1101 CCAATCAAAA AAATCTACG GCATCGGATT CGCCGAAAAA ACCGTTCCCG
1151 TCTGGCACA AGACGTGCGC TATTTGAAT TGCAACAAA CGGCAAAACC
1201 ATCGCGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCG CGCCTTTGCC GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGCCAC GGAAGTGCAC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGGCAT CAACGGCGTA AAA

```

This corresponds to the amino acid sequence <SEQ ID 3094; ORF 128-1.ng>:

g128-1.pep (partial)

```

1  MIDNALLHLG EEPFENQIKT EDIKPAVQTA IAEARGQIAA VKAQHTGTWA
51  NTVRLTGIT ERVGRIWGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDAAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDLSYAGEKL REAKYAFSET
351 EVKKYFPVGK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
401 IGGVYMDLYA REGKRGGAWM NDYKRRRFA DGTLLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILILFHETGH GLHLLTQVD ELGVSGINGV K

```

m128-1/g128-1 ORFs 128-1 and 128-1.ng showed a 94.5% identity in 491 aa overlap

10 20 30 40 50 60

g128-1.pep	MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGTWANTVERLTGIT
m128-1	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTWANTVEPLTGIT
	10 20 30 40 50 60
g128-1.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
m128-1	ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
	70 80 90 100 110 120
g128-1.pep	TLSPAQKTKLDHDLRDFVLSGAELPPERQAEELAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQAEELAKLQTEGAQLSAKFSQNVLDATDAFGIY
	130 140 150 160 170 180
g128-1.pep	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
	190 200 210 220 230 240
g128-1.pep	TRASELSNDGKFDNTANIDRTLENALKTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL
	250 260 270 280 290 300
g128-1.pep	ARRAKPYAEKDLAEVKAFAREHLGLADPQPDLSYAGEKLREAKYAFSETEVKKYFPVGK
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLPQPDLSYAGEKLREAKYAFSETEVKKYFPVGK
	310 320 330 340 350 360
g128-1.pep	VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQONGKTIGGVYMDLYAREGKRGGAWM
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQONGETIGGVYMDLYAREGKRGGAWM
	370 380 390 400 410 420
g128-1.pep	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHGLHLLTQVD
	430 440 450 460 470 480
g128-1.pep	ELGVSGINGVK
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
	490 500 510 520 530 540

The following DNA sequence was identified in *N. meningitidis* <SEQ ID 3095>:

a128-1.seq

```

1  ATGACTGACA ACGCACTGCT CCATTGGGC GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCG GAACTGCGCG
251 CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAATC AACACGATC
401 TGCGCGATT TCGCTCAGC GCGCGGAAC TGCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTCCG CCRAATTCTC

```

```

501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTCATCC AATACGCGCA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCGCGCCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTCGCCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCGC
1101 CCAAATCAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAAA CGGCGAAACC
1201 ATAGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAGAAG CCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAAC TCCAGTCAG
1501 TTTATGGAAA ATTTCTGTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AAATCTCCAA CGCGGAATGT TCCTCGTCCG CCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAGAA GTCGCGCTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCGCCA
1901 CAGGCAACG CTTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 3096; ORF 128-1.a>:

a128-1.pep

```

1 MTDNALLHLG EEPRFDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGW
51 NTVEPLTGIT ERVGRIWGVV SHLNSVDTDP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEKGTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLPQ WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTQLQPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPELKF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRKKNW QQVLDVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNA*

```

m128-1/a128-1 ORFs 128-1 and 128-1.a showed a 97.8% identity in 677 aa overlap

```

          10      20      30      40      50      60
a128-1.pep MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
|||||
m128-1      MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
          10      20      30      40      50      60

          70      80      90      100     110     120
a128-1.pep ERVGRIWGVVSHLNSVDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
|||||:|||||:|||||
m128-1      ERVGRIWGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
          70      80      90      100     110     120

          130     140     150     160     170     180

```

a128-1.pep	TLSHAQKTKLNHDLRDFVL	SGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
m128-1	TLSPAQKTKLNHDLRDFVL	SGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY
	130	140 150 160 170 180
a128-1.pep	190	200 210 220 230 240
	FDDAAPLAGIPEDALAMFAAAQSE	GKTGYKIGLQIPHYLAVIQYADNRKLRQIYRAYV
m128-1	FDDAAPLAGIPEDALAMFAAAQSE	SKTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV
	190	200 210 220 230 240
a128-1.pep	250	260 270 280 290 300
	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMDTPEQVNLFLHDL	
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMDTPEQVNLFLHDL	
	250	260 270 280 290 300
a128-1.pep	310	320 330 340 350 360
	ARRAKPYAEKDLAEVKAFARESLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGK	
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLPWDLGYASEKLREAKYAFSETEVKKYFPVGK	
	310	320 330 340 350 360
a128-1.pep	370	380 390 400 410 420
	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM	
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM	
	370	380 390 400 410 420
a128-1.pep	430	440 450 460 470 480
	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD	
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD	
	430	440 450 460 470 480
a128-1.pep	490	500 510 520 530 540
	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ	
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ	
	490	500 510 520 530 540
a128-1.pep	550	560 570 580 590 600
	RGMFLVRQMEFALFDMMIYSEDDGRLKNWQQVLDVSRKEVAVVRPPEYNRFANSFGHIF	
m128-1	RGMFLVRQMEFALFDMMIYSEDDGRLKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIF	
	550	560 570 580 590 600
a128-1.pep	610	620 630 640 650 660
	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG	
m128-1	AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGGSRSAAESFKAFRG	
	610	620 630 640 650 660
a128-1.pep	670	679
	REPSIDALLRHSGFDNAAX	
m128-1	REPSIDALLRHSGFDNAVX	
	670	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3097>:

m206.seq

m206.pcp..

g206.seq

g206.pep

m206/q206

	10	20	30	40	50	60
m206.pep	MFP	PD	KT	FL	CL	SALL
	10	20	30	40	50	60
g206	MF	SP	DK	TL	FL	CL
	70	80	90	100	110	120
m206.pep	LGL	IG	TP	YK	WG	GS
	70	80	90	100	110	120
g206	LGL	IG	TP	YK	WG	GS
	130	140	150	160	170	
m206.pep	L	V	F	N	T	G
	130	140	150	160	170	
g206	I	V	F	N	T	G

130 140 150 160 170

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3101>:

```
a206.seq
1  ATGTTTCCCC CCGACAAAAC CCTTTTCCTC TGTCTCAGCG CACTGCTCCT
51  CGCCTCATGC GGCACGACCT CCGGCAAACA CCGCCAACCG AAACCCAAAC
101 AGACAGTCCG GCAAATCCAA GCCGTCCGCA TCAGCCACAT CGACCGCACA
151 CAAGGCTCGC AGGAACTCAT GCTCCACAGC CTCGGACTCA TCGGCACGCC
201 CTACAAATGG GCGGGCAGCA GCACCGCAAC CGGCTTCGAT TGCAGCGGCA
251 TGATTCAATT CGTTTACAAA AACGCCCTCA ACGTCAAGCT GCCGCGCACC
301 GCGCGCGACA TGGCGGCGGC AAGCCGCAAA ATCCCCGACA GCCGCCTTAA
351 GCGCGGCGAC CTCGTATTCT TCAACACCGG CGGCGCACAC CGCTACTCAC
401 ACGTCGGACT CTATATCGGC AACGGCGAAT TCATCCATGC CCCCAGCAGC
451 GGCAAAACCA TCAAACCGA AAAACTCTCC ACACCGTTT ACGCCAAAAA
501 CTACCTCGGC GCACATACTT TCTTTACAGA ATGA
```

This corresponds to the amino acid sequence <SEQ ID 3102; ORF 206.a>:

```
a206.pep
1  MFPPDKTLFL CLSALLLASC GTTSGKHRQP KPKQTVRQIQ AVRISHIDRT
51  QGSQELMLHS LGLIGTPYKW GGSSTATGFD CSGMIQFVYK NALNVKLPRT
101 ARDMAAASRK IPDSRLKAGD LVFFNTGGAH RYSHVGLYIG NGEFIHAPSS
151 GKTIKTEKLS TPFYAKNYLG AHTFFTE*
```

m206/a206 ORFs 206 and 206.a showed a 99.4% identity in 177 aa overlap

	10	20	30	40	50	60
m206.pep	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
a206	MFPPDKTLFLCLSALLLASC GTTSGKHRQPKPKQTVRQIQAVRISHIDRTQGSQELMLHS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m206.pep	LGLIGTPYKWGGSSTATGFD CSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRXKAGD					
a206	LGLIGTPYKWGGSSTATGFD CSGMIQFVYKNALNVKLPRTARDMAAASRKIPDSRLKAGD					
	70	80	90	100	110	120
	130	140	150	160	170	
m206.pep	LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX					
a206	LVFFNTGGAHRYSHVGLYIGNGEFIHAPSSGKTIKTEKLSTPFYAKNYLGAHTFFTEX					
	130	140	150	160	170	

287

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3103>:

```
m287.seq
1  ATGTTTAAAC GCAGCGTAAT CGCAATGGCT TGTATTTTTC CCCTTTCAGC
51  CTGCGGGGGC GGCGGTGGCG GATCGCCCGA TGTCAGTCG GCGGACACGC
101 TGTCAAACCC TGCCGCCCTT GTTGTCTCTG AAAAAGAGAC AGAGGCAAAG
151 GAAGATGCGC CACAGGCAGG TTCTCAAGGA CAGGGCGCGC CATCCGCACA
201 AGGCAGTCAA GATATGGCGG CGGTTTCGGA AGAAAATACA GGCAATGGCG
251 GTGCGGTAAC AGCGGATAAT CCCAAAAATG AAGACGAGGT GGCACAAAAT
301 GATATGCCGC AAAATGCCGC CGGTACAGAT AGTTCGACAC CGAATCACAC
351 CCCGGATCCG AATATGCTTG CCGGAAATAT GGAAATCAA GCAACGGATG
401 CCGGGGAATC GTCTCAGCCG GCAAACCAAC CGGATATGGC AAATGCGGCG
451 GACGGAATGC AGGGGGACGA TCCGTCCGCA GGCGGGCAAA ATGCCGGCAA
501 TACGGCTGCC CAAGGTGCAA ATCAAGCCGG AAACAATCAA GCCGCCGGTT
551 CTTAGATCC CATCCCGCG TCAAACCTG CACCTGCGAA TGGCGGTAGC
```

```

601 AATTTTGGAA GGGTTGATTT GGCTAATGGC GTTTTGATTG ACGGGCCGTC
651 GCAAAATATA ACGTTGACCC ACTGTAAAGG CGATTCTTGT AGTGGCAATA
701 ATTTCTTGGA TGAAGAAGTA CAGCTAAAT CAGAATTGA AAAATTAAGT
751 GATGCAGACA AAATAAGTAA TTACAAGAAA GATGGGAAGA ATGATAAATT
801 TGTCGGTTTG GTTGCCGATA GTGTGCAGAT GAAGGGAATC AATCAATATA
851 TTATCTTTTA TAAACCTAAA CCCACTTCAT TTGCGCGATT TAGGCGTTCT
901 GCACGGTCGA GGCGGTCGCT TCCGGCCGAG ATGCCGCTGA TTCCCGTCAA
951 TCAGGCGGAT ACGCTGATTG TCGATGGGGA AGCGGTCAGC CTGACGGGGC
1001 ATTCCGGCAA TATCTTCGCG CCCGAAGGGA ATTACCGGTA TCTGACTTAC
1051 GGGGCGGAAA AATTGCCCGG CGGATCGTAT GCCCTTCGTG TTCAAGGCGA
1101 ACCGGCAAAA GCGGAAATGC TTGCGGGCGC GGCCGTGTAC AACGGCGAAG
1151 TACTGCATTT CCATACGAA AACGGCCGTC CGTACCCGAC CAGGGGCAGG
1201 TTTGCCGCAA AAGTCGATTT CGGCAGCAA TCTGTGGACG GCATTATCGA
1251 CAGCGGCGAT GATTTCGATA TGGGTACGCA AAAATTCAAA GCCGCCATCG
1301 ATGGAACGG CTTTAAGGGG ACTTGACGG AAAATGGCAG CGGGGATGTT
1351 TCCGAAAGT TTACGGCCC GCGCGCGAG GAAGTGGCG GAAAATACAG
1401 CTATCGCCCG ACAGATCGCG AAAAGGGCG ATTCCGCGTG TTTGCCGGCA
1451 AAAAAGAGCA GGATTGA

```

This corresponds to the amino acid sequence <SEQ ID 3104; ORF 287>:

m287.pep

```

1  MEKRSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVSEKETEAK
51  EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT GNGGAVTADN PKNEDEVAQN
101 DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP ANQPDMAAAN
151 DGMQGDPSA GGQNAAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
201 NGRVRLANG VLDGSPQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS
251 DADKISNYKK DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS
301 ARSRRSLPAE MPLIPVNOAD TLIVDGEAVS LTGHSNIFA PEGNYRYLTY
351 GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY NGEVLHFHTE NGRPYPTRGR
401 FAAKVDFGSK SVDGIIDSGD DLHMGTKFKF AAIDGNGFKG TWTENGSGDV
451 SGKIFYGPAGE EVAGKYSYRP TDAEKGFGV FAGKKEQD*

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3105>:

g287.seq

```

1  atgtttaaac gcagtgtgat tgcaatggct tgtatttttc ccctttcagc
51  ctgtgggggc ggcggtggcg gatcgcccga tgtcaagtcg gcggacacgc
101 cgtcaaaacc ggccgcccc gttgttgctg aaaaatgccg ggaaggggtg
151 ctgccgaaaag aaaagaaaga tgaggaggca gcgggcggtg cgccgcaagc
201 cgatacgtag gacgcaaccg ccggagaagg cagccaagat atggcggcag
251 tttcggcaga aaatacaggc aatggcggtg cggcaacaac ggacaacccc
301 aaaaatgaag acgcgggggc gcaaaatgat atgccgcaaa atgccgcccga
351 atccgcaaat caaacaggga acaaccaacc cgccggttct tcagattccg
401 ccccgcgctc aaaccctgcc cctgcgaatg gcggtagcga ttttggaagg
451 acgaacgtgg gcaattctgt tgtgattgac ggaccgtcgc aaaatataac
501 gttgaccacac tgtaaaaggcg attcttgtaa tgggtgataat ttattggatg
551 aagaagcacc gtcaaaatca gaatttgaaa aattaagtga tgaagaaaaa
601 attaaagcat ataaaaaaga cgagcaacgg gagaattttg tcggttttgt
651 tgctgacagg gtaaaaaagg atggaactaa caaatatata atcttctata
701 cggacaaacc acctactcgt tctgcacggt cgaggaggtc gcttccggcc
751 gagattccgc tgattcccgt caatcaggcc gatacgctga ttgtggatgg
801 ggaagcggtc agcctgacgg ggcattccgg caatatcttc gcgccgaag
851 ggaattaccg gtatctgact tacggggcgg aaaaattgcc cgcggtatcg
901 tatgccctcc gtgtgcaagg cgaaccggca aaaggcgaaa tgcttggttg
951 cacggccgtg tacaacggcg aagtgtgca tttccatatg gaaaacggcc
1001 gtccgtaccc gtccggaggc aggtttgccg caaaagtcca tttcggcagc
1051 aaatctgtgg acggcattat cgacagcgcc gatgatttgc atatgggtac
1101 gcaaaaattc aaagccgcca tcgatggaaa cggctttaag gggacttggg
1151 cggaatagg cggcggggat gtttcgggaa ggttttacgg cccggccggc
1201 gaggaagtgg cgggaaaata cagctatcgc ccgacagatg ctgaaaaggg
1251 cggattcggc gtgtttgccg gcaaaaaaga tcgggattga

```

This corresponds to the amino acid sequence <SEQ ID 3106; ORF 287.ng>:

g287.pep

```

1  MEKRSVIAMA CIFPLSACGG GGGGSPDVKS ADTPSKPAAP VVAENAGEGV

```

```

51 LPKEKKDEEA AGGAPQADTQ DATAGEGSQD MAAVSAENTG NGGAATTDNP
101 KNEDAGAQN DMPQNAAESAN QTGNNQPAGS SDSAPASNPA PANGGSDGFR
151 TNVGNVSVID GPSQNTLTH CKGDSCNGDN LLDEEAPSKS EFEKLSDEEK
201 IKRYKKDEQR ENFVGLVADR VKKDGTKNYI IFYTDKPTR SARSRRLPA
251 EIPLIPVNA DTLIVDGEAV SLTGHSGNIF APEGNYRYLT YGAEKLPGGS
301 YALRVQGEPA KGEMLVGTAV YNGEVLHFHM ENGRPYPSGG RFAAKVDFGS
351 KSVDDGIIDSG DDLHMGTKF KAAIDGNGFK GTWTENGGGD VSGRFYGPAG
401 EEVAGKYSYR PTDAEKGFGF VFAGKKDRD*

```

m287/g287 ORFs 287 and 287.ng showed a 70.1% identity in 499 aa overlap

```

m287.pep      10      20      30      40      49
MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPAPVSE-----KETE
|||||
g287          10      20      30      40      50      60
MFKRSVIAMACIFPLSACGGGGGSPDVKSADTPSKPAPVVAENAGEGVLPEKKDEEA

m287.pep      50      60      70      80      90     100     109
KEDAPQAGSQGQGPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNMPQNAAGT
||||:| |::|
g287          70      80      90     100     110
AGGAPQADTQD--ATAGEGSQDMAAVSAENTGNGGAATTDNPKNEDAGAQNMPQNA--

m287.pep     110     120     130     140     150     160     169
DSSTPNHTPDNMLAGNMENQATDAGESSQFANQPDMANAADGMQDDPSAGGQNAGNTA
g287          -----

m287.pep     170     180     190     200     210     220     229
AQQANQAGNNQAAGSSDPIPASNPAPANGGSNFGVRDLANGVLIDGPSQNTLTHCKGDS
::|:|
g287         120     130     140     150     160     170
-ESANQTGNNQPAGSSDSAPASNPAANGGSDFGRTNVGNVSVIDGPSQNTLTHCKGDS

m287.pep     230     240     250     260     270     280     289
CSGNNFLDEEVQLKSEFEKLSADKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKP
|:|:|
g287         180     190     200     210     220     230
CNGDNLLEDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFVGLVADRVRKKGDTNKYIIFYTD

m287.pep     290     300     310     320     330     340     349
KPTSFARFRRSARSRSLPAEMPLIPVNAQDTLIVDGEAVSLTGHSGNIFAPEGNYRYLT
||:
g287         240     250     260     270     280     290
KPPT-----RSARSRSLPAEIPLIPVNAQDTLIVDGEAVSLTGHSGNIFAPEGNYRYLT

m287.pep     350     360     370     380     390     400     409
YGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGS
|||||
g287         300     310     320     330     340     350
YGAEKLPGGSYALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPYPSGGRFPAKVDFGS

m287.pep     410     420     430     440     450     460     469
KSVDDGIIDSGDDLHMGTKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYR
|||||
g287         360     370     380     390     400     410
KSVDDGIIDSGDDLHMGTKFKAAIDGNGFKGTWTENGGGDVSGRFYGPAGEEVAGKYSYR

m287.pep     470     480     489
PTDAEKGFGFVFAGKKEQDX

```


g287 |||||||:|
 PTDAEKGFGVFAGKKDRDX
 420 430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3107>:

a287.seq
 1 ATGTTTAAAC GCAGTGTGAT TGCAATGGCT TGTATTGTTG CCCTTTCAGC
 51 CTGTGGGGGC GCGGTGGCG GATCGCCCGA TGTTAAGTCG GCGGACACGC
 101 TGTCAAAACC TGCCGCCCCCT GTTGTACTG AAGATGTCGG GGAAGAGGTG
 151 CTGCCGAAAG AAAAGAAAGA TGAGGAGGCG GTGAGTGGTG CGCCGCAAGC
 201 CGATACGCAG GACGCAACCG CCGGAAAAGG CCGTCAAGAT ATGGCGGCAG
 251 TTTCCGCAGA AAATACAGGC AATGGCGGTG CCGCAACAAC GGATAATCCC
 301 GAAAAATAAG ACGAGGGACC GCAAATGAT ATGCCGCAAA ATGCCGCCGA
 351 TACAGATAGT TCGACACCGA ATCACACCCC TGCACCGAAT ATGCCAACCA
 401 GAGATATGGG AAACCAAGCA CCGGATGCCG GGAATCGGC ACAACCGGCA
 451 AACCAACCGG ATATGGCAAA TCGGCGGCAC GGAATGCAGG GGGACGATCC
 501 GTCGGCAGGG GAAATGCCC GCAATACGGC AGATCAAGCT GCAAATCAAG
 551 CTGAAACAA TCAAGTCGGC GGCTCTCAA ATCCTGCCTC TTCAACCAAT
 601 CCTAACGCCA CGAATGGCGG CAGCGATTTT GGAAGGATAA ATGTAGCTAA
 651 TGGCATCAAG CTTGACAGCG GTTCGAAAA TGTAACGTTG ACACATTGTA
 701 AAGACAAAGT ATGCGATAGA GATTTCTTAG ATGAAGAAGC ACCACCAAAA
 751 TCAGAATTTG AAAAATTAAG TGATGAAGAA AAAATTAATA AATATAAAAA
 801 AGACGAGCAA CGAGAGAATT TTGTCGGTTT GGTGCTGAC AGGGTAGAAA
 851 AGAATGGAAC TAACAAATAT GTCATCATT ATAAAGACAA GTCCGCTTCA
 901 TCTTCATCTG CGCGATTCAG GCGTCTGCA CCGTCGAGGC GGTGCTTCC
 951 GGCCGAGATG CCGCTGATTC CCGTCAATCA GCGGATACG CTGATTGTCG
 1001 ATGGGGAAGC GGTACGCTG ACGGGGCATT CCGGCAATAT CTTGCGGCCC
 1051 GAAGGGAATT ACCGGTATCT GACTTACGGG GCGGAAAAAT TGTCCGCGCG
 1101 ATCGTATGCC CTCAGTGTGC AAGGCGAACC GGCAAAAGGC GAAATGCTTG
 1151 CGGGCACGGC CGTGTACAAC GCGGAAGTGC TGCATTTCCA TATGGAAAC
 1201 GGCCGTCCGT CCGGTCCGG AGGCAGGTTT GCGCAAAAG TCGATTTCGG
 1251 CAGCAAATCT GTGGACGGCA TTATCGACAG CCGCGATGAT TTGCATATGG
 1301 GTACGCAAAA ATTCAAAGCC GTTATCGATG GAAACGGCTT TAAGGGGACT
 1351 TGGACGGAAT ATGGCGGCGG GGATGTTTCC GGAAGGTTT ACGCCCGGC
 1401 CCGCGAAGAA GTGGCGGGAA AATACAGCTA TCGCCCGACA GATCGGAAA
 1451 AGGCGGGATT CGGCGTGTTC GCCGCAAAA AAGAGCAGGA TTGA

This corresponds to the amino acid sequence <SEQ ID 3108; ORF 287.a>:

a287.pep
 1 MFKRSVIAMA CIVALSAACG GGGGSPDVKS ADTLSKPAAP VVTEDVGEEV
 51 LPKEKKDEEA VSGAPQADTQ DATAGKGGQD MAAVSAENTG NGGAATDNP
 101 ENKDEGPQND MPQNAADTDS STPNHTPAPN MPTRDMGNQA PDAGESAQPA
 151 NQPDMANAAD GMQDDPSAG ENAGNTADQA ANQAENNQVG GSQNPASSTN
 201 PNATNGGSDF GRINVANGIK LDSGSENVTL THCKDKVCDR DFLDEEAPPK
 251 SEFEKLSDEE KINKYKKDEQ RENFVGLVAD RVEKNGTNKY VVIYKDKSAS
 301 SSSARFRRSA RSRRLPAEM PLIPVNQADT LIVDGEAVSL TGHSGNIFAP
 351 EGNRYRLTYG AEKLSGGSYA LSVQGEPAKG EMLAGTAVYN GEVLHFHMFN
 401 GRPSPSGGRF AAKVDFGSKS VDGIIDSGDD LHMGTQKFKA VIDGNGFKGT
 451 WTENGGGDVS GRFYGPAGEE VAGKYSYRPT DAEKGGFGVF AGKKEQD*

m287/a287 ORFs 287 and 287.a showed a 77.2% identity in 501 aa overlap

	10	20	30	40	49
m287.pep	MFKRSVIAMACIFALSACGGGGGGSPDVKSADTLSKPAAPVVSE-----	KETEA			
a287	MFKRSVIAMACIVALSACGGGGGGSPDVKSADTLSKPAAPVVTE	DVGEEVLPKEKKDEEA			
	10	20	30	40	50
	60	70	80	90	100
m287.pep	KEDAPQAGSQGGAPSAQGSQDMAAVSEENTGNGGAVTADNP	KNEDEVAQNDMQNAAGT			
a287	VSGAPQADTQ--DATAGKGGQDMAAVSAENTGNGGAATT	DNPENKDEGPQNDMPQNAADT			
	70	80	90	100	110

```

m287.pep      110      120      130      140      150      160      169
               DSSTPNHTPDNMLAGNMENQATDAGESSQPANQPDMANAADMGMQDDPSAGGQONAGNTA
a287          110      120      130      140      150      160      169
               DSSTPNHTPDNMPTRDMGNQAPDAGESAQPANQPDMANAADMGMQDDPSAG-ENAGNTA
               120      130      140      150      160      170

m287.pep      170      180      190      200      210      220      229
               AQGANQAGNNQAAGSSDPIPASNPAPANGGSNFGVRVLANGVLIDGPSQNTLTHCKGDS
a287          170      180      190      200      210      220      229
               DQAANQAENNQVGGSONPASSTNPATNGGSDFGGRINVANGIKLDSGSENVTLTHCKDKV
               180      190      200      210      220      230

m287.pep      230      240      250      260      270      280      289
               CSGNNFLDEEVQLKSEFEKLSADAKISNYKKDGKNDKFVGLVADSVQMKGINQYIIFYKP
a287          230      240      250      260      270      280      289
               CD-RDFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFVGLVADRVKNGTKYVYIYKD
               240      250      260      270      280      290

m287.pep      290      300      310      320      330      340
               KP--TSFARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRY
a287          290      300      310      320      330      340
               KSASSSSARFRRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRY
               300      310      320      330      340      350

m287.pep      350      360      370      380      390      400
               LTYGAEKLPGGSYALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGREPYPTRGRFAAKVDF
a287          350      360      370      380      390      400
               LTYGAEKLSGGSYALSQGEPAKGEMLAGTAVYNGEVLHFHMENGRFSPSGGRFAAKVDF
               360      370      380      390      400      410

m287.pep      410      420      430      440      450      460
               GSKSVDDGIIDSGDDLHMGTKQFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVACKYS
a287          410      420      430      440      450      460
               GSKSVDDGIIDSGDDLHMGTKQFKAVIDGNGFKGTWTENGSGDVSGRFYGPAGEEVACKYS
               420      430      440      450      460      470

m287.pep      470      480      489
               YRPTDAEKGFGVFAGKKEQDX
a287          470      480      489
               YRPTDAEKGFGVFAGKKEQDX
               480      490

```

406

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3109>:

m406.seq

```

1  ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG ACCTGACAG GTATTCCATC GCATGGCGGA GGTAACGCT
101 TTGCGGTCTGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 CACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301 GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CTCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451 ATGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501 CGACACTGCC TTCTTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTTCTCCT GCCAATGCCG ATACAGATGT GTTTATTAAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAACCT GGAATATTTC GCAGTAGACA

```

```

701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACGGAAG GATTAATGGT CGATTCTCC GATATCCGAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GTAGTGCAC AACATAGACA
951 AGGACAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3110; ORF 406>:

m406.pep

```

1  MQARLLIPIIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSGLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSAGLTG LTSLSTLNA PALSRTQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPKVSKGIK PTEGLMVDFS DIRPYGNHTG NSAPSVEADN
301 SHEGYGYSDE VVRQHRQGP *

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 3111>:

g406.seq

```

1  ATGCGGGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG AACTGACAG GTATTCCATC GCATGGCGGA GGCAAACGCT
101 TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTATAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGTACTCCA
251 TTGATGCACT GATTCGCGGC GAATACATAA ACAGCCCTGC CGTCCGCACC
301 GATTACACCT ATCCGCGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACGGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCCTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA GGAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CCAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTGCA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTTCTCCTT GCCAATGCCG ATACAGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAACT GGAATATTTT GCAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGGCCGTATA AAGTAAGCAA
801 AGGAATCAAA CCGACGGAAG GATTGATGGT CGATTCTCC GATATCCAAC
851 CATACGGCAA TCATACGGGT AACTCCGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCAC AACATAGACA
951 AGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3112; ORF 406>:

g406.pep

```

1  MRARLLIPIIL FSVFILSACG TLTGIPSHGG GKRFAVEQEL VAASARAAVK
51  DMDLQALHGR KVALYIATMG DQSGSGLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPRYETT AETTSAGLTG LTSLSTLNA PALSRTQSDG SGSRSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVP ANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA
251 AYKENYALWM GPKVSKGIK PTEGLMVDFS DIQPYGNHTG NSAPSVEADN
301 SHEGYGYSDE AVRQHRQGP *

```

ORF 406 shows 98.8% identity over a 320 aa overlap with a predicted ORF (ORF406.a) from *N. gonorrhoeae*:

g406/m406

	10	20	30	40	50	60
g406.pep	MRARLLIPIILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA	AVK	DMDLQALHGR			
m406	MQARLLIPIILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA	AVK	DMDLQALHGR			

110

	10	20	30	40	50	60
	70	80	90	100	110	120
g406.pep	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
m406	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTSGGLTG					
	130	140	150	160	170	180
g406.pep	LTTSLSLTINAPALSRTQSDGSGSRSSLGLNIGMGDYRNETLTNPRDTAFLSHLVQTVF					
m406	LTTSLSLTINAPALSRTQSDGSGSKSSLGLNIGMGDYRNETLTNPRDTAFLSHLVQTVF					
	190	200	210	220	230	240
g406.pep	FLRGIDVVS PANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
m406	FLRGIDVVS PANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	250	260	270	280	290	300
g406.pep	IKPKTNAFEAAAYKENYALWMGPYKVS KG IKPTEGLMVD FSDIQPYGNHTGNSAPSVEADN					
m406	IKPKTNAFEAAAYKENYALWMGPYKVS KG IKPTEGLMVD FSDIRPYGNHTGNSAPSVEADN					
	310	320				
g406.pep	SHEGYGYSDEAVRQHRQGQPX					
m406	SHEGYGYSDEVVRQHRQGQPX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3113>:

```

a406.seq
1   ATGCAAGCAC GGCTGCTGAT ACCTATTCTT TTTTCAGTTT TTATTTTATC
51  CGCCTGCGGG AACTGACAG GTATTCCATC GCATGGCGGA GGTAACGCT
101 TCGCGGTCGA ACAAGAACTT GTGGCCGCTT CTGCCAGAGC TGCCGTAAAA
151 GACATGGATT TACAGGCATT ACACGGACGA AAAGTTGCAT TGTACATTGC
201 AACTATGGGC GACCAAGGTT CAGGCAGTTT GACAGGGGGT CGCTACTCCA
251 TTGATGCACT GATTCGTGGC GAATACATAA ACAGCCCTGC CGTCCGTACC
301 GATTACACCT ATCCACGTTA CGAAACCACC GCTGAAACAA CATCAGGCGG
351 TTTGACAGGT TTAACCACTT CTTTATCTAC ACTTAATGCC CCTGCACTCT
401 CGCGCACCCA ATCAGACGGT AGCGGAAGTA AAAGCAGTCT GGGCTTAAAT
451 ATTGGCGGGA TGGGGGATTA TCGAAATGAA ACCTTGACGA CTAACCCGCG
501 CGACACTGCC TTTCTTTCCC ACTTGGTACA GACCGTATTT TTCCTGCGCG
551 GCATAGACGT TGTCTCTCCT GCCAATGCCG ATACGGATGT GTTTATTAAC
601 ATCGACGTAT TCGGAACGAT ACGCAACAGA ACCGAAATGC ACCTATACAA
651 TGCCGAAACA CTGAAAGCCC AAACAAAAC TGAATATTTC CGAGTAGACA
701 GAACCAATAA AAAATTGCTC ATCAAACCAA AAACCAATGC GTTTGAAGCT
751 GCCTATAAAG AAAATTACGC ATTGTGGATG GGACCGTATA AAGTAAGCAA
801 AGGAATTAAA CCGACAGAAG GATTAATGGT CGATTTCTCC GATATCCAAC
851 CATACGGCAA TCATATGGGT AACTCTGCCC CATCCGTAGA GGCTGATAAC
901 AGTCATGAGG GGTATGGATA CAGCGATGAA GCAGTGCGAC GACATAGACA
951 AGGGCAACCT TGA

```

This corresponds to the amino acid sequence <SEQ ID 3114; ORF 406.a>:

```

a406.pep
1   MQARLLIPIL FSVFILSACG TLTGIPSHGG GKRFVEQEL VAASARAQVK
51  DMDLQALHGR KVALYIATMG DQSGSGLTGG RYSIDALIRG EYINSPAVRT
101 DYTYPYRYET AETTSGGLTG LTTSLSLTNA PALSRTQSDG SGSKSSLGLN
151 IGGMGDYRNE TLTTNPRDTA FLSHLVQTVF FLRGIDVVS PANADTDVFIN
201 IDVFGTIRNR TEMHLYNAET LKAQTKLEYF AVDRTNKKLL IKPKTNAFEA

```

251 AYKENYALWM GPYKVSKGIK PTEGLMVDFS DIQPYGNHMG NSAPSVEADN
 301 SHEGYGYSDE AVRRHRQGQP *

m406/a406 ORFs 406 and 406.a showed a 98.8% identity in 320 aa overlap

	10	20	30	40	50	60
m406.pep	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA AVKMDLQALHGR					
a406	MQARLLIPILFSVFILSACGTLTGIPSHGGGKRFAVEQELVAASARA AVKMDLQALHGR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m406.pep	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTS GGLTG					
a406	KVALYIATMGDQSGSLTGGRYSIDALIRGEYINSPAVRTDYTPRYETTAETTS GGLTG					
	70	80	90	100	110	120
	130	140	150	160	170	180
m406.pep	LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGMGDYRNETLT TNPRDTAFLSHLVQTVF					
a406	LTTSLSTLNAPALSRTQSDGSGSKSSLGLNIGMGDYRNETLT TNPRDTAFLSHLVQTVF					
	130	140	150	160	170	180
	190	200	210	220	230	240
m406.pep	FLRGIDVVS PANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
a406	FLRGIDVVS PANADTDVFINIDVFGTIRNRTEMHLYNAETLKAQTKLEYFAVDRTNKKLL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m406.pep	IKPKTNAFEAA YKENYALWMGPYKVSKGIKPTEGLMVDFS DIRPYGNHTGNSAPSVEADN					
a406	IKPKTNAFEAA YKENYALWMGPYKVSKGIKPTEGLMVDFS DIQPYGNHMGNSAPSVEADN					
	250	260	270	280	290	300
	310	320				
m406.pep	SHEGYGYSDEVVRQHRQGQPX					
a406	SHEGYGYSDEAVRRHRQGQPX					
	310	320				

EXAMPLE 2

Expression of ORF 919

The primer described in Table 1 for ORF 919 was used to locate and clone ORF 919. The predicted gene 919 was cloned in pET vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 919-His fusion protein purification. Mice were immunized with the purified 919-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; PP, purified protein, TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm

that 919 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 919 are provided in Figure 10. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 919 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 3

Expression of ORF 279

The primer described in Table 1 for ORF 279 was used to locate and clone ORF 279. The predicted gene 279 was cloned in pGex vector and expressed in *E. coli*. The product of protein expression and purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 279-GST purification. Mice were immunized with the purified 279-GST and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 279 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 279 are provided in Figure 11. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 279 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 4

Expression of ORF 576 and 576-1

The primer described in Table 1 for ORF 576 was used to locate and clone ORF 576. The predicted gene 576 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 576-GST fusion protein purification. Mice were immunized with the purified 576-GST and sera

were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that ORF 576 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 576 are provided in Figure 12. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 576 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 5

Expression of ORF 519 and 519-1

The primer described in Table 1 for ORF 519 was used to locate and clone ORF 519. The predicted gene 519 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 519-His fusion protein purification. Mice were immunized with the purified 519-His and sera were used for Western blot (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 519 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 519 are provided in Figure 13. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 519 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 6

Expression of ORF 121 and 121-1

The primer described in Table 1 for ORF 121 was used to locate and clone ORF 121. The predicted gene *121* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 121-His fusion protein purification. Mice were immunized with the purified 121-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 121 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 121 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 121 are provided in Figure 14. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 121 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 7

Expression of ORF 128 and 128-1

The primer described in Table 1 for ORF 128 was used to locate and clone ORF 128. The predicted gene *128* was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 128-His purification. Mice were immunized with the purified 128-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D) and ELISA assay (panel E). Results show that 128 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 128 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipathic regions of ORF 128 are provided in Figure 15. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J.*

Immunol 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 128 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 8

Expression of ORF 206

The primer described in Table 1 for ORF 206 was used to locate and clone ORF 206. The predicted gene 206 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 206-His purification. Mice were immunized with the purified 206-His and sera were used for Western blot analysis (panel B). It is worthnoting that the immunoreactive band in protein extracts from meningococcus is 38 kDa instead of 17 kDa (panel A). To gain information on the nature of this antibody staining we expressed ORF 206 in *E. coli* without the His-tag and including the predicted leader peptide. Western blot analysis on total protein extracts from *E. coli* expressing this native form of the 206 protein showed a recative band at a position of 38 kDa, as observed in meningococcus. We conclude that the 38 kDa band in panel B) is specific and that anti-206 antibodies, likely recognize a multimeric protein complex. In panel C is shown the FACS analysis, in panel D the bactericidal assay, and in panel E) the ELISA assay. Results show that 206 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 206 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 519 are provided in Figure 16. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 206 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 9

Expression of ORF 287

The primer described in Table 1 for ORF 287 was used to locate and clone ORF 287. The predicted gene 287 was cloned in pGex vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 287-GST fusion protein purification. Mice were immunized with the purified 287-GST and sera were used for FACS analysis (panel B), bactericidal assay (panel C), and ELISA assay (panel D). Results show that 287 is a surface-exposed protein. Symbols: M1, molecular weight marker. Arrow indicates the position of the main recombinant protein product (A). These experiments confirm that 287 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 287 are provided in Figure 17. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al. 1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 287 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 10

Expression of ORF 406

The primer described in Table 1 for ORF 406 was used to locate and clone ORF 406. The predicted gene 406 was cloned in pET vector and expressed in *E. coli*. The product of protein purification was analyzed by SDS-PAGE. In panel A) is shown the analysis of 406-His fusion protein purification. Mice were immunized with the purified 406-His and sera were used for Western blot analysis (panel B), FACS analysis (panel C), bactericidal assay (panel D), and ELISA assay (panel E). Results show that 406 is a surface-exposed protein. Symbols: M1, molecular weight marker; TP, *N. meningitidis* total protein extract; OMV, *N. meningitidis* outer membrane vesicle preparation. Arrows indicate the position of the main recombinant protein product (A) and the *N. meningitidis* immunoreactive band (B). These experiments confirm that 406 is a surface-exposed protein and that it is a useful immunogen. The hydrophilicity plots, antigenic index, and amphipatic regions of ORF 406 are provided in Figure 18. The AMPHI program is used to predict putative T-cell epitopes (Gao et al 1989, *J. Immunol* 143:3007; Roberts et al. 1996, *AIDS Res Human Retroviruses* 12:593; Quakyi et al.

1992, *Scand J Immunol Suppl* 11:9). The nucleic acid sequence of ORF 406 and the amino acid sequence encoded thereby is provided in Example 1.

EXAMPLE 11

Table 2 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 225 among different strains.

Table 2

225 gene variability: List of used <i>Neisseria</i> strains	
Identification Strains number	Source / reference
Group B	
zo01_225 NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo02_225 BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zo03_225 NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zo04_225 297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zo05_225 1000	R. Moxon / Seiler <i>et al.</i> , 1996
zo06_225 BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zo07_225 BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zo08_225 528	R. Moxon / Seiler <i>et al.</i> , 1996
zo09_225 NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zo10_225 BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zo11_225 NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zo12_225 NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zo13_225 NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zo14_225 NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zo15_225 SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zo16_225 NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zo17_225 NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zo18_225 BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zo19_225 BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zo20_225 44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zo21_225 MC58	R. Moxon
zo96_225 2996	Our collection
Group A	
zo22_225 205900	R. Moxon
zo23_225 F6124	R. Moxon
z2491 Z2491	R. Moxon / Maiden <i>et al.</i> , 1998
Group C	
zo24_225 90/18311	R. Moxon
zo25_225 93/4286	R. Moxon

Others

zo26_225 A22 (group W) R. Moxon / Maiden *et al.*, 1998
 zo27_225 E26 (group X) R. Moxon / Maiden *et al.*, 1998
 zo28_225 860800 (group Y) R. Moxon / Maiden *et al.*, 1998
 zo29_225 E32 (group Z) R. Moxon / Maiden *et al.*, 1998

Gonococcus

zo32_225 Ng F62 R. Moxon / Maiden *et al.*, 1998
 zo33_225 Ng SN4 R. Moxon

fa1090 FA1090 R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
 Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

>FA1090 <SEQ ID 3115>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGSAMGLLGIAIYRYGGTSVSTGFDCS
 GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
 RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKNDPSRFLN*

Z2491 <SEQ ID 3116>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRVPARRAGNA
 DELIGNAMGLNEQPVLPVNRAPARRAGNADELIGNAMGLLGIAIYRYGGTSISTGFDCSGF
 MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRF
 IHAPRTGKNIEITSLSHKYWSGKYAFARRVKNDPSRFLN*

Z001_225 <SEQ ID 3117>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
 DELIGNAMGLLGIAIYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
 VKNDPSRFLN*

Z002_225 <SEQ ID 3118>
 MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
 DELIGNAMGLLGIAIYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
 VKNDPSRFLN*

Z003_225 <SEQ ID 3119>
 MDSFFKPAVWAVLWLMFAVRLALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
 DELIGNAMGLLGIAIYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIAPRTGKNIEITSLSHKYWSGKYAFARR
 VKNDPSRFLN*

Z004_225 <SEQ ID 3120>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z005_225 <SEQ ID 3121>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGSAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z006_225 <SEQ ID 3122>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z007_225 <SEQ ID 3123>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z008_225 <SEQ ID 3124>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGSAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z009_225 <SEQ ID 3125>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z010_225 <SEQ ID 3126>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z011_225 <SEQ ID 3127>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLNEQPVL PVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRFI
HAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z012_225 <SEQ ID 3128>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVL PINRAPARRAG
NADELIGSAMGLNEQPVL PVNRVPARRAGNADELIGNAMGLNEQPVL PVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z013_225 <SEQ ID 3129>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPVNRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFIQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z014_225 <SEQ ID 3130>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z015_225 <SEQ ID 3131>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAHPTGKNIEITSLSHKYWSGKYAFARRVKKNDSRFLN*

Z016_225 <SEQ ID 3132>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z017_225 <SEQ ID 3133>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z018_225 <SEQ ID 3134>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z019_225 <SEQ ID 3135>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z020_225 <SEQ ID 3136>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPINRAPARRAGNADELIGSAMGLNEQPVLVPVNRVPARRAGNA
DELIGNAMGLNEQPVLVPVNRAPARRAGNADELIGNAMGLLGIAYRYGGTSVSTGFDCSGF
MQHIFKRAMGINLPRTSAEQARMGTPVARSELQPGDMVFFRTLGGSRISHVGLYIGNNRFI
HAPRTGKNIEITSLSHKYWSGKYAFARRVKKNDSRFLN*

Z021_225 <SEQ ID 3137>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKNDSRFLN*

Z022_225 <SEQ ID 3138>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLVPINRAPARRAG
NADELIGSAMGLNEQPVLVPVNRVPARRAGNADELIGNAMGLNEQPVLVPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR

SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z023_225 <SEQ ID 3139>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z024_225 <SEQ ID 3140>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z025_225 <SEQ ID 3141>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z026_225 <SEQ ID 3142>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z027_225 <SEQ ID 3143>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z028_225 <SEQ ID 3144>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z029_225 <SEQ ID 3145>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
NADELIGSAMGLNEQPVLPVNRVPARRAGNADELIGNAMGLNEQPVLPVNRAPARRAGNA
DELIGNAMGLLGIAYRYGGTSVSTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
SELQPGDMVFFRTLGGSRISHVGLYIGNNRFIHAPRTGKNIEITSLSHKYWSGKYAFARR
VKKN DPSRFLN*

Z032_225 <SEQ ID 3146>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z033_225 <SEQ ID 3147>

MDSFFKPAVWAVLWLMFAVRSALADELTNLLSSREQILRQFAEDEQPVLPVNRAPARRAG
NADELIGSAMGLNEQPVLPVNRAPARRAGNADELIGSAMGLLGIAYRYGGTSVSTGFDCS
GFMQHIFKRAMGINLPRTSAEQARMGAPVARSELQPGDMVFFRTLGGSRISHVGLYIGNN
RFIHAPRTGKNIEITSLSHKYWSGKYAFARRVKKN DPSRFLN*

Z096_225 <SEQ ID 3148>

MDSFFKPAVWAVLWLMFAVRPALADELTNLLSSREQILRQFAEDEQPVLPINRAPARRAG
 NADELIGSAMGLNEQPVLVNRVPARRAGNADELIGNAMGLNEQPVLVNRAPARRAGNA
 DELIGNAMGLLGIAYRYGGTSISTGFDCSGFMQHIFKRAMGINLPRTSAEQARMGTPVAR
 SELQPGDMVFFRTLGGSRISHVGLYIGNNRFTHAPRTGKNIEITSLSHKYWSGKYAFARR
 VKKNDPSRFLN*

Figure 19 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 225, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 12

Table 3 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 235 among different strains.

Table 3

235 gene variability: List of used *Neisseria* strains

Identification Strains number	Reference
Group B	
gnmzq01 NG6/88	Seiler <i>et al.</i> , 1996
gnmzq02 BZ198	Seiler <i>et al.</i> , 1996
gnmzq03 NG3/88	Seiler <i>et al.</i> , 1996
gnmzq04 1000	Seiler <i>et al.</i> , 1996
gnmzq05 1000	Seiler <i>et al.</i> , 1996
gnmzq07 BZ169	Seiler <i>et al.</i> , 1996
gnmzq08 528	Seiler <i>et al.</i> , 1996
gnmzq09 NGP165	Seiler <i>et al.</i> , 1996
gnmzq10 BZ133	Seiler <i>et al.</i> , 1996
gnmzq11 NGE31	Seiler <i>et al.</i> , 1996
gnmzq13 NGE28	Seiler <i>et al.</i> , 1996
gnmzq14 NGH38	Seiler <i>et al.</i> , 1996
gnmzq15 SWZ107	Seiler <i>et al.</i> , 1996
gnmzq16 NGH15	Seiler <i>et al.</i> , 1996
gnmzq17 NGH36	Seiler <i>et al.</i> , 1996
gnmzq18 BZ232	Seiler <i>et al.</i> , 1996
gnmzq19 BZ83	Seiler <i>et al.</i> , 1996
gnmzq21 MC58	Virji <i>et al.</i> , 1992
Group A	
gnmzq22 205900	Our collection

gnmzq23	F6124	Our collection
z2491	Z2491	Maiden <i>et al.</i> , 1998

Group C

gnmzq24	90/18311	Our collection
gnmzq25	93/4286	Our collection

Others

gnmzq26	A22	(group W) Maiden <i>et al.</i> , 1998
gnmzq27	E26	(group X) Maiden <i>et al.</i> , 1998
gnmzq28	860800	(group Y) Maiden <i>et al.</i> , 1998
gnmzq29	E32	(group Z) Maiden <i>et al.</i> , 1998
gnmzq31		<i>N. lactamica</i> Our collection

Gonococcus

gnmzq32	Ng F62	Maiden <i>et al.</i> , 1998
gnmzq33	Ng SN4	Our collection

fa1090	FA1090	Dempsey <i>et al.</i> 1991
--------	--------	----------------------------

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
 Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.
 Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279
 Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3149>

MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
 AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVGAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

GNMZQ01 <SEQ ID 3150>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANNL
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ02 <SEQ ID 3151>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ03 <SEQ ID 3152>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST

AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ04 <SEQ ID 3153>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ05 <SEQ ID 3154>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ07 <SEQ ID 3155>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ08 <SEQ ID 3156>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ09 <SEQ ID 3157>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ10 <SEQ ID 3158>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ11 <SEQ ID 3159>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ13 <SEQ ID 3160>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ14 <SEQ ID 3161>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ15 <SEQ ID 3162>
MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGT
YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ16 <SEQ ID 3163>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ17 <SEQ ID 3164>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ18 <SEQ ID 3165>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ19 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ21 <SEQ ID 3166>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ22 <SEQ ID 3167>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ23 <SEQ ID 3168>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ24 <SEQ ID 3169>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ25 <SEQ ID 3170>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ26 <SEQ ID 3171>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ27 <SEQ ID 3172>

MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
YQILDSVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ28 <SEQ ID 3173>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ29 <SEQ ID 3174>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

GNMZQ31 <SEQ ID 3175>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
 AEPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITITEYGTS
 YQILDVTTVSARARLVDSRNGKVLWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
 DRGYQVSKAAAYDLLSPYSHNGILKGPRFVEEQPK*

GNMZQ32 <SEQ ID 3176>
 MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
 AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

GNMZQ33 <SEQ ID 3177>
 MKPLILGLAAVLALSACQVRKAPDLDYTSFKESKPASILVVPPLNESPDVNGTWGMLAST
 AAPISEAGYYVFPAAVVEETFKENGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVGAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSRNGILKGPRFVEEQPK*

Z2491 <SEQ ID 3178>
 MKPLILGLAAVLALSACQVQKAPDFDYTSFKESKPASILVVPPLNESPDVNGTWGVLAST
 AAPLSEAGYYVFPAAVVEETFKQNGLTNAADIHAVRPEKLHQIFGNDVLYITVTEYGTS
 YQILDVTTVSAKARLVDSRNGKELWSGSASIREGSNNSNSGLLGALVSAVVNQIANSLT
 DRGYQVSKTAAYNLLSPYSHNGILKGPRFVEEQPK*

Figure 20 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 235, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 13

Table 4 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 287 among different strains.

Table 4

287 gene variability: List of used *Neisseria* strains

Identification Strains number	Reference
-------------------------------	-----------

Group B

287_2	BZ198	Seiler <i>et al.</i> , 1996
287_9	NGP165	Seiler <i>et al.</i> , 1996
287_14	NGH38	Seiler <i>et al.</i> , 1996
287_21	MC58	Virji <i>et al.</i> , 1992

Group A

z2491	Z2491	Maiden <i>et al.</i> , 1998
-------	-------	-----------------------------

Gonococcus

fa1090	FA1090	Dempsey <i>et al.</i> 1991
--------	--------	----------------------------

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
 Maiden R. *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.
 Virji M. *et al.*, Mol. Microbiol., 1992, 6:1271-1279
 Dempsey J.F. *et al.*, J. Bacteriol., 1991, 173:5476-5486

The amino acid sequences for each listed strain are as follows:

287_14 <SEQ ID 3179>

MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
 QGAPSAQGGQDMAAVSEENTGNGGAAATDKPKNEDEGAQNMPQNAADTDSLTPNHTPAS
 NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNQ
 TAGSQNPASSTNPSATNSGGDFGRNTVGNVVIDGPSQNTLTHCKGDCSGNNFLDEEV
 QLKSEFEKLSADAKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPSTSFAR
 FRRSARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLP
 GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII
 DSGDGLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKG
 GFGVFAGKKEQD*

287_2 <SEQ ID 3180>

MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
 QGAPSAQGGQDMAAVSEENTGNGGAAATDKPKNEDEGAQNMPQNAADTDSLTPNHTPAS
 NMPAGNMENQAPDAGESEQPANQPDMAANTADGMQGGDDPSAGGENAGNTAAQGTNQAENNQ
 TAGSQNPASSTNPSATNSGGDFGRNTVGNVVIDGPSQNTLTHCKGDCSGNNFLDEEV
 QLKSEFEKLSADAKISNYKKDGKNDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPSTSFAR
 FRRSARSRRLPAEMPLIPVNQADTLIVDGEAVSLTGHSNIFAPEGNYRYLTYGAEKLP
 GGSYALRVQGEPSKGEMLAGTAVYNGEVLHFHTENGRPSPSRGRFAAKVDFGSKSVDGII
 DSGDGLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKG
 GFGVFAGKKEQD*

287_21. <SEQ ID 3181>

MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
 QGAPSAQGSQDMAAVSEENTGNGGAVTADNPKNEDEVAQNMPQNAAGTDSSTPNHTPDP
 NMLAGNMENQATDAGESSQPANQPDMAANAADGMQGGDDPSAGGQNAAGNTAAQGANQAGNNQ
 AAGSSDPIPASNPAPANGSNFGRVDLANGVLIDGPSQNTLTHCKGDCSGNNFLDEEV
 QLKSEFEKLSADAKISNYKKDGKNDKFVGLVADSVQMKGINQYIIIFYKPKPSTSFARFRRS

ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVVDGIIDSGD
DLHMGTKQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGFGV
FAGKKEQD*

287_9 <SEQ ID 3182>

MFKRSVIAMACIVALSACGGGGGSPDVKSADTLSPKPAAPVVTEDEVGEEVLPKEKKDEEA
VSGAPQADTQDATAGKGGQDMAAVSAENTGNNGAATTNPNKDEGPQNDMPQNAADTDS
STPNHTPAPNMPTRDMGNQAPDAGESAQPANQPDMANAADMGGDDPSAGENAGNTADQA
ANQAENNOVGGSQNPASSTNPATNGGSDFRINVANGIKLDSGSENVTLTHCKDKVCDR
DFLDEEAPPKSEFEKLSDEEKINKYKKDEQRENFGVLVADRVEKNGTNKYVIYKDKSAS
SSSARFRSARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYG
AEKLSGGSYALSVQGEPAKGEMLAGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKS
VDGIIDSGDDLHMGTKQKFAVIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPT
DAEKGGFGVFAGKKEQD*

FA1090 <SEQ ID 3183>

MFKRSVIAMACIFPLSACGGGGGSPDVKSADTPSKPAAPVVAENAGEGVLPKEKKDEEA
AGGAPQADTQDATAGESSQDMAAVSAENTGNNGAATTNPNKEDAGAQNMPQNAAESAN
QTGNNQFAGSSDSAPASNPAFANGGSDFRITNVGNSVVIDGPSQNTLTHCKGDSNGDN
LLDEEAPSKSEFEKLSDEEKIKRYKKDEQRENFGVLVADRVEKNGTNKYIIFYTDKPPTR
SARSRRSLPAEIPPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGS
YALRVQGEPAKGEMLVGTAVYNGEVLHFHMENGRPSPSGGRFAAKVDFGSKSVVDGIIDSG
DDLHMGTKQKFAAIDGNGFKGTWTENGSGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGFG
VFAGKKDRD*

Z2491 <SEQ ID 3184>

MFKRSVIAMACIFALSACGGGGGSPDVKSADTLSPKPAAPVVSEKETEAKEDAPQAGSQG
QGAPSAQGSQDMAAVSEENTGNNGAVTADNPKNEDEVAQNMPQNAAGTDSSTPNHTPDP
NMLAGNMENQATDAGESSQPANQPDMANAADMGGDDPSAGGQNAAGTAAQGANQAGNNQ
AAGSSDPIPASNPAPANGGNSNFRVLDLANGVLIDGPSQNTLTHCKGDSGNNFLDEEV
QLKSEFEKLSADAKISNYKDKGNDKFVGLVADSVQMGKINQYIIFYKPKPFSFARFRS
ARSRRSLPAEMPLIPVNQADTLIVDGEAVSLTGHSGNIFAPEGNYRYLTYGAEKLPGGSY
ALRVQGEPAKGEMLAGAAVYNGEVLHFHTENGRPYPTRGRFAAKVDFGSKSVVDGIIDSGD
DLHMGTKQKFKAAIDGNGFKGTWTENGSGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGFGV
FAGKKEQD*

Figure 21 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 287, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 14

Table 5 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 519 among different strains.

Table 5

519 gene variability: List of used *Neisseria* strains

Identification Strains	Source / reference
------------------------	--------------------

number**Group B**

zv01_519	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv02_519	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zv03_519ass	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zv04_519	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zv05_519	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zv06_519ass	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zv07_519	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zv11_519	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zv12_519	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zv18_519	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zv19_519	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zv20_519ass	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zv21_519ass	MC58	R. Moxon
zv96_519	2996	Our collection

Group A

zv22_519ass	205900	R. Moxon
z2491_519	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

Others

zv26_519	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zv27_519	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zv28_519	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zv29_519ass	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998

Gonococcus

zv32_519	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
fa1090_519	FA1090	R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
 Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090_519 <SEQ ID 3185>
 MEFFIILLAAVAVFGFKSFVVIPQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHS
 KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
 RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRAMQAQITAERE
 KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQA AVNASNAEKIARINRAKGEAESLR
 LVAEANAETIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
 ISAGMKIIDSSKTAK*

Z2491_519 <SEQ ID 3186>

MEFFIILLAAVVVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV01_519 <SEQ ID 3187>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV02_519 <SEQ ID 3188>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV03_519 <SEQ ID 3189>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV04_519 <SEQ ID 3190>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV05_519 <SEQ ID 3191>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV06_519ASS <SEQ ID 3192>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV07_519 <SEQ ID 3193>

MEFFIILLVAVAVFGFKSFVVI PQQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAEAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV11_519 <SEQ ID 3194>

MEFFIILLAAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV12_519 <SEQ ID 3195>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV18_519 <SEQ ID 3196>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV19_519 <SEQ ID 3197>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV20_519ASS <SEQ ID 3198>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSM
ISAGMKIIDSSKTAK*

ZV21_519ASS <SEQ ID 3199>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV22_519ASS <SEQ ID 3200>

MEFFIILLAAVVVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVSALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV26_519 <SEQ ID 3201>

MEFFIILLAAVVVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV27_519 <SEQ ID 3202>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEIILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV28_519 <SEQ ID 3203>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEIILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV29_519ASS <SEQ ID 3204>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEIILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV32_519 <SEQ ID 3205>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEIILRAMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

ZV96_519 <SEQ ID 3206>

MEFFIILLVAVAVFGFKSFVVIPOQEVHVVERLGRFHRALTAGLNILIPFIDRVAYRHSL
KEIPLDVPSQVCITRDNTQLTVDGIIYFQVTDPKLASYGSSNYIMAITQLAQTTLRSVIG
RMELDKTFEERDEINSTVVAALDEAAGAWGVKVLRYEIKDLVPPQEIILRSMQAQITAERE
KRARIAESEGRKIEQINLASGQREAEIQQSEGEAQAAVNASNAEKIARINRAKGEAESLR
LVAEANAIAIRQIAAALQTQGGADAVNLKIAEQYVAAFNNLAKESNTLIMPANVADIGSL
ISAGMKIIDSSKTAK*

Figure 22 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 519, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 15

Table 6 lists several *Neisseria* strains which were used to assess the conservation of the sequence of ORF 919 among different strains.

Table 6

919 gene variability: List of used *Neisseria* strains

Identification Strains	Source / reference
------------------------	--------------------

number**Group B**

zm01	NG6/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm02	BZ198	R. Moxon / Seiler <i>et al.</i> , 1996
zm03	NG3/88	R. Moxon / Seiler <i>et al.</i> , 1996
zm04	297-0	R. Moxon / Seiler <i>et al.</i> , 1996
zm05	1000	R. Moxon / Seiler <i>et al.</i> , 1996
zm06	BZ147	R. Moxon / Seiler <i>et al.</i> , 1996
zm07	BZ169	R. Moxon / Seiler <i>et al.</i> , 1996
zm08n	528	R. Moxon / Seiler <i>et al.</i> , 1996
zm09	NGP165	R. Moxon / Seiler <i>et al.</i> , 1996
zm10	BZ133	R. Moxon / Seiler <i>et al.</i> , 1996
zm11asbc	NGE31	R. Moxon / Seiler <i>et al.</i> , 1996
zm12	NGF26	R. Moxon / Seiler <i>et al.</i> , 1996
zm13	NGE28	R. Moxon / Seiler <i>et al.</i> , 1996
zm14	NGH38	R. Moxon / Seiler <i>et al.</i> , 1996
zm15	SWZ107	R. Moxon / Seiler <i>et al.</i> , 1996
zm16	NGH15	R. Moxon / Seiler <i>et al.</i> , 1996
zm17	NGH36	R. Moxon / Seiler <i>et al.</i> , 1996
zm18	BZ232	R. Moxon / Seiler <i>et al.</i> , 1996
zm19	BZ83	R. Moxon / Seiler <i>et al.</i> , 1996
zm20	44/76	R. Moxon / Seiler <i>et al.</i> , 1996
zm21	MC58	R. Moxon
zm96	2996	Our collection

Group A

zm22	205900	R. Moxon
zm23asbc	F6124	R. Moxon
z2491	Z2491	R. Moxon / Maiden <i>et al.</i> , 1998

Group C

zm24	90/18311	R. Moxon
zm25	93/4286	R. Moxon

Others

zm26	A22	(group W) R. Moxon / Maiden <i>et al.</i> , 1998
zm27bc	E26	(group X) R. Moxon / Maiden <i>et al.</i> , 1998
zm28	860800	(group Y) R. Moxon / Maiden <i>et al.</i> , 1998
zm29asbc	E32	(group Z) R. Moxon / Maiden <i>et al.</i> , 1998
zm31asbc	<i>N. lactamica</i>	R. Moxon

Gonococcus

zm32asbc	Ng F62	R. Moxon / Maiden <i>et al.</i> , 1998
zm33asbc	Ng SN4	R. Moxon
fa1090	FA1090	R. Moxon

References:

Seiler A. *et al.*, Mol. Microbiol., 1996, 19(4):841-856.
Maiden *et al.*, Proc. Natl. Acad. Sci. USA, 1998, 95:3140-3145.

The amino acid sequences for each listed strain are as follows:

FA1090 <SEQ ID 3207>
MKKHLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSMPHWAADFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRTERARFPIYGIIPDDFISVPLPAGLRGGKN
LVRIQRTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGVPVAGLGTPLMGEYAGA
IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

Z2491 <SEQ ID 3208>
MKKYLFRALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKOFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELTSNDGVPVAGLGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM01 <SEQ ID 3209>
MKKYLFRALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKOFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMQRNPQRLAEVLGQNPSYIFFRELAGSSNDGVPVAGLGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM02 <SEQ ID 3210>
MKKYLFRALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKOFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMQRNPQRLAEVLGQNPSYIFFRELAGSSNDGVPVAGLGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM03 <SEQ ID 3211>
MKKYLFRALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKOFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMQRNPQRLAEVLGQNPSYIFFRELAGSSNDGVPVAGLGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM04 <SEQ ID 3212>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGQTSMQGIKAYMQONPQRLAEVLGQNPSYIFFRELTSNDGSPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM05 <SEQ ID 3213>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D P A G T T V G G G G A V
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGSPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM06 <SEQ ID 3214>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D P A G T T V G G G G A V
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGSPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM07 <SEQ ID 3215>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D P A G T T V G G G G A V
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGSPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM08N <SEQ ID 3216>

MKKYLFRAALYGI A A A I L A A C Q S K S I Q T F P Q P D T S V I N G P D R P V G I P D P A G T T V G G G G A V
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGSPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM09 <SEQ ID 3217>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELTSNDGSPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM10 <SEQ ID 3218>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEPEVLKGGDDRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL

KLQGTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM11ASBC <SEQ ID 3219>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGKYMADKGYL
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM12 <SEQ ID 3220>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM13 <SEQ ID 3221>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLQGTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM14 <SEQ ID 3222>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGKYMADKGYL
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSRNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM15 <SEQ ID 3223>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDLAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNHQQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGKYMADKGYL
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSGNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM16 <SEQ ID 3224>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPGRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYGIIPDDFISVPLPAGLRSGKA
LVRIRQTGKNSGTIDNTGGHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGKYMADKGYL
KLQGTSMQGIKSYMRQNPQRLAEVLGQNPSYIFFRELTGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM17 <SEQ ID 3225>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM18 <SEQ ID 3226>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM19 <SEQ ID 3227>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM20 <SEQ ID 3228>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM21 <SEQ ID 3229>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSRFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM22 <SEQ ID 3230>

MKKYLFRAALCGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
KLGQTSMQGIKAYMQNPORLAEVLGQNPSYIFFRELTSNDGPVGALGTPLMGEYAGA
VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM23ASBC <SEQ ID 3231>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVPVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYEYEPVLKGDDRRTAQARFPIYIGIPDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIOGSGRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL

KLGTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELAGSSNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK
MKEPGYVWQLLPNGMKPEYRP*

ZM24 <SEQ ID 3232>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGKYMADKGYL
KLGTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM25 <SEQ ID 3233>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPAPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGKYMADKGYL
KLGTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM26 <SEQ ID 3234>

MKKYLFRAALYGI AAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSVQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGTSMQGIKAYMQNPORLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM27BC <SEQ ID 3235>

MKKYLFRAALYGISAAAILAACQSKSIQTFFQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNAGGHTADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGETAGK
MKEPGYVWQLLPNGMKPEYRP*

ZM28 <SEQ ID 3236>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGTSMQGIKAYMRQNPORLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM29ASBC <SEQ ID 3237>

MKKYLFRAALCGIAAAILAACQSKSIQTFFQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
YFTPWQVAGNGSLAGTVTGYYEPVLKGDDRRTAQARFPIYGIPTDDFISVPLPAGLRSGKA
LVRIQRTGKNSGTIDNTGGTHTADLSQFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVVSIGRYMADKGYL
KLGTSMQGIKSYMRQNPORLAEVLGQNPSYIFFRELTSNGNDGPVGALGTPLMGEYAGA
VDRHYITLGAFLFVATHTPITRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
QKTTGYVWQLLPNGMKPEYRP*

ZM31ASBC <SEQ ID 3238>

MKKHLFRAALYGIAAAILAACQSKSIQTFPQPDTSIIKGPDRPAGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM32ASBC <SEQ ID 3239>

MKKHLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
 YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYGI PDDFISVPLPAGLRGGKA
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGDGPVAGLGTPLMGGYAGA
 IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM33ASBC <SEQ ID 3240>

MKKHLRSALYGIAAAILAACQSRSIQTFPQPDTSVINGPDRPAGIPDPAGTTVAGGGAV
 YTVVPHLSMPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKRFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRTERARFPIYGI PDDFISVPLPAGLRGGKN
 LVRI RQTGKNSGTIDNAGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSGNEGPGVAGLGTPLMGEYAGA
 IDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

ZM96 <SEQ ID 3241>

MKKYLFRAALYGIAAAILAACQSKSIQTFPQPDTSVINGPDRPVGIPDPAGTTVGGGGAV
 YTVVPHLSLPHWAAQDFAKSLQSFRLGCANLKNRQGWQDVCAQAFQTPVHSFQAKQFFER
 YFTPWQVAGNGSLAGTVTGYEYEPVLKGDGRRRTAQARFPIYGI PDDFISVPLPAGLRSGKA
 LVRI RQTGKNSGTIDNTGGTHADLSRFPITARTTAIKGRFEGSRFLPYHTRNQINGGAL
 DGKAPILGYAEDPVELFFMHIIQSGSRLKTPSGKYIRIGYADKNEHPYVSIGRYMADKGYL
 KLGQTSMQGIKAYMRQNPQRLAEVLGQNPSYIFFRELAGSSNDGPGVAGLGTPLMGEYAGA
 VDRHYITLGAPLFVATAHPVTRKALNRLIMAQDTGSAIKGAVRVDYFWGYGDEAGELAGK
 QKTTGYVWQLLPNGMKPEYRP*

Figure 23 shows the results of aligning the sequences of each of these strains. Dark shading indicates regions of homology, and gray shading indicates the conservation of amino acids with similar characteristics. As is readily discernible, there is significant conservation among the various strains of ORF 919, further confirming its utility as an antigen for both vaccines and diagnostics.

EXAMPLE 16

Using the above-described procedures, the following oligonucleotide primers were employed in the polymerase chain reaction (PCR) assay in order to clone the ORFs as indicated:

Table 7: Oligonucleotides used for PCR to amplify complete or partial ORFs

ORF primer Sequence		Restriction sites
001 Forward	CGCGGATCCCATATG-TGGATGGTGCTGGTCAT	BamHI- NdeI
Reverse	CCCGCTCGAG-TGCCGTCTTGTCAC	XhoI
003 Forward	CGCGGATCCCATATG-GTCGTATTCGTGGC	BamHI- NdeI
Reverse	CCCGCTCGAG-AAAATCATGAACACGCGC	XhoI
005 Forward	CGCGGATCCCATATG-GACAATATTGACATGT	BamHI- NdeI
Reverse	CCCGCTCGAG-CATCACATCCGCCCCG	XhoI
006 Forward	CGCGGATCCCATATG-CTGCTGGTGCTGG	BamHI- NdeI
Reverse	CCCGCTCGAG-AGTTCCGGCTTTGATGT	XhoI
007 Forward	CGCGGATCCCATATG-GCCGACAACAGCATCAT	BamHI- NdeI
Reverse	CCCGCTCGAG-AAGGCGTTCATGATATAAG	XhoI
008 Forward	CGCGGATCCCATATG-AACAACAGACATTTTG	BamHI- NdeI
Reverse	CCCGCTCGAG-CCTGTCCGGTAAAAGAC	XhoI
009 Forward	CGCGGATCCCATATG-CCCCGCGCTGCT	BamHI- NdeI
Reverse	CCCGCTCGAG-TGGCTTTTGCCACGTTTT	XhoI
011 Forward	CGCGGATCCCATATG-AAGACACACCGCAAG	BamHI- NdeI
Reverse	CCCGCTCGAG-GGCGGTCAGTACGGT	XhoI
012 Forward	CGCGGATCCCATATG-CTCGCCCGTTGCC	BamHI- NdeI
Reverse	CCCGCTCGAG-AGCGGGGAAGAGGCAC	XhoI
013 Forward	CGCGGATCCCATATG-CCTTTGACCATGCT	BamHI- NdeI
Reverse	CCCGCTCGAG-CTGATTCGGCAAAAAAATCT	XhoI
018 Forward	CGCGGATCCCATATG-CAGCAGAGGCAGTT	BamHI- NdeI
Reverse	CCCGCTCGAG-GACGAGGCGAACGCC	XhoI
019 Forward	AAAGAATTC-CTGCCAGCCGGCAAGACCCCGGC	Eco RI
Reverse	AAACTGCAG-TCAGCGGGCGGGGACAATGCCCAT	Pst I
023 Forward	AAAGAATTC-AAAGAATATTCGGCATGGCAGGC	Eco RI
Reverse	AAACTGCAG-TTACCCCCAAATCACTTTAACTGA	Pst I
025 Forward	AAAGAATTC-TGCGCCACCCAACAGCCTGCTCC	Eco RI
Reverse	AAACTGCAG-TCAGAACGCGATATAGCTGTTCGG	Pst I
031 Forward	CGCGGATCCCATATG-GTCTCCCTTCGCTT	BamHI- NdeI
Reverse	CCCGCTCGAG-ATGTAAGACGGGGACAAC	XhoI
032 Forward	CGCGGATCCCATATG-CGGCGAAACGTGC	BamHI-

	Reverse	CCCGCTCGAG-CTGGTTTTTTGATATTTGTG	NdeI
033	Forward	CGCGGATCCCATATG-GCGGCGGCAGACA	XhoI
	Reverse	CCCGCTCGAG-ATTTGCCGCATCCCGAT	BamHI-
034	Forward	CGCGGATCCCATATG-GCCGAAAACAGCTACGG	NdeI
	Reverse	CCCGCTCGAG-TTTGACGATTTGGTTCAATT	XhoI
036	Forward	CGCGGATCCCATATG-CTGAAGCCGTGCG	BamHI-
	Reverse	CCCGCTCGAG-CCGGACTGCGTATCGG	NdeI
038	Forward	CGCGGATCCCATATG-ACCGATTTCGCCA	XhoI
	Reverse	CCCGCTCGAG-TTCTACGCCGTACTGCC	BamHI-
039	Forward	CGCGGATCCCATATG-CCGTCCGAACCGC	NdeI
	Reverse	CCCGCTCGAG-TAGGATGACGAGGTAGG	XhoI
041	Forward	CGCGGATCCCATATG-TTCGTGCGCAACCGC	BamHI-
	Reverse	CCCGCTCGAG-GCCCAAAAACCTCTTTCAA	NdeI
042	Forward	CGCGGATCCCATATG-ACGATGATTTGCTTGC	XhoI
	Reverse	CCCGCTCGAG-TTTGCAGCCTGCATTTGAC	BamHI-
043	Forward	AAAAAAGGTACC-ATGGTTGTTCAAATCAAATATC	NdeI
	Reverse	AAACTGCAG-TTATTGCGCTTCACCTTCCGCCGC	XhoI
043a	Forward	AAAAAAGGTACC-GCAAAAGTGCATGGCGGCTTGGACGGTGC	Kpn I
	Reverse	AAAAAACTGCAG-	Pst I
		TTAATCCTGCAACACGAATTCGCCCGTCCG	
044	Forward	CGCGGATCCCATATG-CCGTCCGACTAGAG	Kpn I
	Reverse	CCCGCTCGAG-ATGCGCTACGGTAGCCA	Pst I
046	Forward	AAAGAATTC-ATGTCGGCAATGCTCCCGACAAG	BamHI-
	Reverse	AAACTGCAG-TCACTCGGCGACCCACACCGTGAA	NdeI
047	Forward	CGCGGATCCCATATG-GTCATCATAACAGGCG	XhoI
	Reverse	CCCGCTCGAG-TCCGAAAAAGCCCATTTTG	Eco RI
048	Forward	AAAGAATTC-ATGCTCAACAAAGGCGAAGAATTGCC	NdeI
	Reverse	AAACTGCAG-TCAAGATTGACGGGGATGATGCC	XhoI
049	Forward	AAAGAATTC-ATGCGGGCGCAGGCGTTTGATCAGCC	Eco RI
	Reverse	AAACTGCAG-AAGGCGTATCTGAAAAAATGGCAG	Pst I
050	Forward	CGCGGATCCCATATG-GGCGCGGGCTGG	BamHI-
	Reverse	CCCGCTCGAG-AATCGGGCCATCTTCGA	NdeI
052	Forward	AAAAAAGAATTC-ATGGCTTTGGTGCGGAGGAAAC	XhoI
	Reverse	AAAAAAGTCGAC-TCAGGCGGCGTTTTTACCTTCCT	Eco RI
052a	Forward	AAAAAAGAATTC-GTGGCGGAGGAAACGGAATATCCGC	Sal I
			Eco RI

	Reverse	AAAAAACTGCAG-TTAGCTGTTTTTGAAACGCCGTCCAACCC	Pst I
073	Forward	CGCGGATCCCATATG-TGTATGCCATATAAGAT	BamHI-
	Reverse	CCCGCTCGAG-CACCGGATTGTCCGAC	NdeI
075	Forward	CGCGGATCCCATATG-CCGTCTTACTTCATC	XhoI
	Reverse	CCCGCTCGAG-ATCACCAATGCCGATTATTT	BamHI-
077a	Forward	AAAAAAGAATTC-GGCGGCATTTTCATCGACACCTTCCT	NdeI
	Reverse	AAAAAACTGCAG-TCAGACGAACATCTGCACAAACGCAAT	XhoI
080	Forward	AAAGAATTC-GCGTCCGGGCTGGTTTGGTTTACAATTC	Eco RI
	Reverse	AAACTGCAG-CTATTCTTCGGATTCTTTTTTCGGG	Pst I
081	Forward	AAAGAATTC-ATGAAACCACTGGACCTAAATTTTCATCTG	Eco RI
	Reverse	AAACTGCAG-TCACTTATCCTCCAATGCCTC	Pst I
082	Forward	AAAGAATTC-ATGTGGTTGTTGAAGTTGCCTGC	Eco RI
	Reverse	AAACTGCAG-TTACGCGGATTCGGCAGTTGG	Pst I
084	Forward	AAAGAATTC-TATCACCCAGAATATGAATACGGCTACCG	Eco RI
	Reverse	AAACTGCAG-TTATACTTGGGCGCAACATGA	Pst I
085	Forward	CGCGGATCCCATATG-GGTAAAGGGCAGGACT	BamHI-
	Reverse	CCCGCTCGAG-CAAAGCCTTAAACGCTTCG	NdeI
086	Forward	AAAAAAGGTACC-TATTTGGCATCAAAAGAAGGCGG	XhoI
	Reverse	AAACTGCAG-TTACTCCACCCGATAACCGCG	Kpn I
087	Forward	AAAGAATTC-ATGGGCGGTAAAACCTTTATGC	Pst I
	Reverse	AAACTGCAG-TTACGCCGCACACGCAATCGC	Eco RI
087a	Forward	AAAAAAGAATTC-AAGCTATTAGGCGTGCCGATTGTGATTCA	Pst I
	Reverse	AAAAAACTGCAG-TTACGCCTGCAAGATGCCAGCTTGCC	Eco RI
088	Forward	AAAAAAGAATTC-ATGTTTTTATGGCTCGCACATTTCAG	Pst I
	Reverse	AAAAAACTGCAG-TCAGCGGATTTTGAGGGTACTCAAACC	BamHI-
089	Forward	CGCGGATCCCATATG-CCGCCCAAATCAC	NdeI
	Reverse	CCCGCTCGAG-TGCGCATACCAAAGCCA	XhoI
090	Forward	CGCGGATCCCATATG-CGCATAGTCGAGCA	BamHI-
	Reverse	CCCGCTCGAG-AGCAAAACGGCGGTACG	NdeI
091	Forward	AAAGAATTC-ATGGAAATACCCGTACCGCCGAGTCC	XhoI
	Reverse	AAACTGCAG-TCAGCGCAGGGGGTAGCCCAAGCC	Eco RI
092	Forward	AAAGAATTC-ATGTTTTTTATTTCAATCCG	Pst I
	Reverse	AAACTGCAG-TCAAATCTGTTTCGACAATGC	Eco RI
093	Forward	AAAGAATTC-ATGCAGAATTTTGCCAAAGTGCC	Pst I
	Reverse	AAACTGCAG-CTATGGCTCGTCATACCGGGC	Eco RI
094	Forward	AAAGAATTC-ATGCCGTACGGAAGCGCATCAACTC	Pst I
	Reverse	AAACTGCAG-TTATCCCGGCCATACCGCCGAACA	Eco RI
095	Forward	AAAGAATTC-ATGTCCTTTCATTTGAACATGGACGG	Pst I
	Reverse	AAACTGCAG-TCAACGCCCGCAGGCACTAACGCCC	Eco RI
096	Forward	AAAGAATTC-ATGGCTCGTCATACCGGGCAGGG	Pst I
			Eco RI

	Reverse AA <u>ACTGCAG</u> -TCAAAGGAAAAGGCCGTCTGAAAAGCG	Pst I
097	Forward AAAGAATTC-ATGGACACTTCAAAACAAACACTGTTG	Eco RI
	Reverse AA <u>ACTGCAG</u> -TCAGCCCAAATACCAGAATTTTCAG	Pst I
098	Forward AAAGAATTC-GATGAACGCAGCCAGCATGGATACG	Eco RI
	Reverse AA <u>ACTGCAG</u> -TTACGACATTCTGATTTGGCA	Pst I
102	Forward AAAAAAGAATTC-GGCCTGATGATTTTGGAAAGTCAACAC	Eco RI
	Reverse AAAAA <u>ACTGCAG</u> -TTATCCTTTAAATACGGGGACGAGTTC	Pst I
105	Forward CGCGGATCCCATATG-TCCGCAAACGAATACG	BamHI- NdeI XhoI
	Reverse CCCGCTCGAG-GTGTTCTGCCAGTTTCAG	Eco RI
107	Forward AAAAAAGAATTC- CTGATGATTTTGGAAAGTCAACACCCATTATCC	
	Reverse AAAAA <u>ACTGCAG</u> -TTATCCTTTAAATACGGGGACGAGTTC	Pst I
107b	Forward AAAAAAGAATTC- GATACCCAAGCCCCCGCCGGCACAACTACTG	Eco RI
	Reverse AAAAA <u>ACTGCAG</u> - TTACGCGTCGCCTTTAAAGTATTTGAGCAGGCTGGAGAC	Pst I
108	Forward AAAGAATTC-ATGTTGCCGGGCTTCAACCG	Eco RI
	Reverse AA <u>ACTGCAG</u> -TTAGCGGTACAGGTGTTTGAAGCA	Pst I
108a	Forward AAAAAAGAATTC-GGTAACACATTCCGGCAGCTTAGACGGTGG	Eco RI
	Reverse AA <u>ACTGCAG</u> -TTAGCGGTACAGGTGTTTGAAGCA	Pst I
109	Forward AAAGAATTC-ATGTATTATCGCCGGGTATGGG	Eco RI
	Reverse AA <u>ACTGCAG</u> -CTAGCCCAAAGATTTGAAGTGTTT	Pst I
111	Forward CGCGGATCCCATATG-TGTTTCGGAACAAACCGC	BamHI- NdeI XhoI
	Reverse CCCGCTCGAG-GCGGAGCAGTTTTTCAAA	
114	Forward CGCGGATCCCATATG-GCTTCCATCACTTCGC	BamHI- NdeI XhoI
	Reverse CCCGCTCGAG-CATCCGCGAAATCGTC	
117	Forward AAAAAAGGTACC-ATGGTCGAAGAACTGGAAGTCTG	Kpn I
	Reverse AA <u>ACTGCAG</u> -TTAAAGCCGGGTAACGCTCAATAC	Pst I
118	Forward AAAGTCGACATGTGTGAGTTCAAGGATATTATAAG	Sal I
	Reverse AAAGCATGC-CTATTTTTTGTGTGAATAATCAAATC	Sph I
121	Forward CGCGGATCCCATATG-GAAACACAGCTTTACAT	BamHI- NdeI XhoI
	Reverse CCCGCTCGAG-ATAATAATATCCCGCGCCC	
122	Forward CGCGGATCCCATATG-GTCATGATTAAAATCCGCA	BamHI- NdeI XhoI
	Reverse CCCGCTCGAG-AATCTTGGTAGATTGGATTT	
125	Forward AAAGAATTC-ATGTCGGGCAATGCCTCCTCTCC	Eco RI
	Reverse AA <u>ACTGCAG</u> -TCACGCCGTTTCAAGACG	Pst I
125a	Forward AAAAAAGAATTC-ACGGCAGGCAGCACCGCCGCACAGGTTTC	Eco RI
	Reverse AAAAA <u>ACTGCAG</u> - TTATTTTGCCACGTCGGTTTCTCCGGTGAACAACGC	Pst I
126	Forward CGCGGATCCCATATG-CCGTCTGAAACCC	BamHI-

	Reverse	CCCGCTCGAG-ATATTCCGCCGAATGCC	NdeI
127	Forward	AAAGAATTC-ATGGAAATATGGAATATGTTGGACACTTG	XhoI
	Reverse	AAACTGCAG-TTAAAGTGTTCGGAGCCGGC	Eco RI
			Pst I
127a	Forward	AAAAAAGAATTC-AAGGAACTGATTATGTGTCTGTCTGGG	Eco RI
	Reverse	AAACTGCAG-TTAAAGTGTTCGGAGCCGGC	Pst I
128	Forward	CGCGGATCCCATATG-ACTGACAACGCACT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GACCGCGTTGTCTGAAA	XhoI
130	Forward	CGCGGATCCCATATG-AAACAACTCCGCGA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-GAATTTTGCACCGGATTG	XhoI
132	Forward	AAAGAATTC-ATGGAACCCTTCAAACCTTAATTTG	Eco RI
	Reverse	AAAAAACTGCAG-TCACCATGTCTGGCATTGAAAAAC	Pst I
134	Forward	CGCGGATCCCATATG-TCCCAAGAAATCCTC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-CAGTTTGACCGAATGTTC	XhoI
135	Forward	CGCGGATCCCATATG-AAATACAAAAGAATCGTATT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAATTCGGTCAGAAGCAGG	XhoI
137	Forward	AAAAAAGGTACC-ATGATTACCCATCCCAATTCGATCC	Kpn I
	Reverse	AAAAAACTGCAG-TCAGTGCTGTTTTTTCATGCCGAA	Pst I
137a	Forward	AAAAAAGAATTC-GGCCGCAAACACGGCATCGGCTTCCT	Eco RI
	Reverse	AAAAAACTGCAG-TTAAGCGGGATGACGCGGCAGCATAACC	Pst I
138	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
141	Forward	AAAGAATTC-ATGAGCTTCAAACCCGATGCCGAAATCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAACAAGCCGTGAATCACGCC	Pst I
142	Forward	CGCGGATCCCATATG-CGTGCCGATTTCATG	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAACTGCTGCACATGGG	XhoI
143	Forward	AAAAAAGAATTC- ATGCTCAGTTTCGGCTTTCTCGGCGTTCAGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCCCGCCGTGTGTTCTTTAAT	Pst I
144	Forward	AAAAAAGAATTC-GGTCTGATCGACGGGCGTGCCGTAAC	Eco RI
	Reverse	AAAAAATCTAGA-TCGGCATCGGCCGGCATATGTCCG	Xba I
146	Forward	AAAAAAGAATTC- CGCCAAGTCGTCATTGACCACGACAAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAGGCATCGGCAAATAGGAACTGGG	Pst I
147	Forward	AAAAAAGAATTC-ACTGAGCAATCGGTGGATTGGAAC	Eco RI
	Reverse	AAAAAATCTAGA-TTAGGTAAAGCTGCGGCCCATTTGCGG	Xba I
148	Forward	AAAAAAGAATTC- ATGGCGTTAAAAACATCAAACCTGGAACACGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGCCCTTCATACAGCCTTCGTTTTG	Xba I
149	Forward	CGCGGATCCCATATG-CTGCTTGACAACAAAGT	BamHI-

	Reverse	CCCGCTCGAG-AAACTTCACGTTACGCC	NdeI
150	Forward	CGCGGATCCCATATG-CAGAACACAAATCCG	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-ATAAACATCACGCTGATAGC	NdeI
151	Forward	AAAAAAGAATTC- ATGAAACAAATCCGCAACATCGCCATCATCGC	XhoI
			Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAGCTTTTAAAGTGCGGCG	Pst I
152	Forward	AAAAAAGAATTC- ATGAAAAACAAAACCAAAGTCTGGGACCTCCC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGGACAGGAGCAGGATGGCGGC	Pst I
153	Forward	AAAAAAGAATTC-ATGGCGTTTGCTTACGGTATGAC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTCATGTTTTTCCGTTTCATT	Pst I
153a	Forward	AAAAAAGAATTC-CGGACTTCGGTATCGGTTCCCCAGCATTG	Eco RI
	Reverse	AAAAAACTGCAG- TTACGCCGACGAAATACTCAGACTTTTCGG	Pst I
154	Forward	CGCGGATCCCATATG-ACTGACAACAGCCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TCGGCTTCCTTTCGGG	XhoI
155	Forward	AAAAAAGAATTC-ATGAAAATCGGTATCCCACGCGAGTC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCCTTTCTTAAACATATTCAGCAT	Pst I
156	Forward	AAAAAAGAATTC-GCACAGCAAACGGTTTTGAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCAGCCGCGACAAACAGCCC	Pst I
157	Forward	CGCGGATCCCATATG-AGGAACGAGGAAAAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAAACACAATATCCCCGC	XhoI
158	Forward	AAAAAAGAATTC-GCGGAGCAGTTGGCGATGGCAAATTCTGC	Eco RI
	Reverse	AAAAAATCTAGA-TTATCCACAGAGATTGTTTCCAGTTC	Xba I
160	Forward	CGCGGATCCCATATG-GACATTCTGGACAAAC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTTTGCCCGCCTTCTTT	XhoI
163	Forward	AAAAAAGGTACC-ACCGTGCCGGATCAGGTGCAGATGTG	Kpn I
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATCCACCTGCTCGTG	Xba I
163a	Forward	AAAAAAGAATTC-CGGCTGGTGCAGATAATGAGCCAGAC	Eco RI
	Reverse	AAAAAATCTAGA-TTACTCTGCCAATCCACCTGCTCGTG	Xba I
164	Forward	CGCGGATCCCATATG-AACCGGACTTATGCC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTTGTTTCCGTCAAACCTGC	XhoI
165	Forward	CGCGGATCCGCTAGC-GCTGAAGCGACAGACG	BamHI-
			NheI
	Reverse	CCCGCTCGAG-AATATCCAATACTTTTCGCG	XhoI
206	Forward	CGCGGATCCCATATG-AAACACCGCCAACCGA	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-TTCTGTAAAAAAAGTATGTGC	XhoI
209	Forward	CGCGGATCCCATATG-CTGCGGCATTTAGGA	BamHI-
			NdeI

	Reverse	CCCGCTCGAG-TACCCCTGAAGGCAAC	XhoI
211	Forward	AAAAAAGAATTC-ATGTTGCGGGTTGCTGCTGC	Eco RI
	Reverse	AAAAAACTGCAG-CTATCCTGCGGATTGGCATTGAAA	Pst I
212	Forward	CGCGGATCCCATATG-GACAATCTCGTATGG	BamHI-
	Reverse	CCCGCTCGAG-AGGGGTTAGATCCTTCC	NdeI
215	Forward	CGCGGATCCCATATG-GCATGGTTGGGTCGT	XhoI
	Reverse	CCCGCTCGAG-CATATCTTTTGTATCATAAATC	BamHI-
216	Forward	CGCGGATCCCATATG-GCAATGGCAGAAAACG	NdeI
	Reverse	CCCGCTCGAG-TACAATCCGTGCCGCC	XhoI
217	Forward	CGCGGATCCCATATG-GCGGATGACGGTGTG	BamHI-
	Reverse	CCCGCTCGAG-ACCCCGAATATCGAATCC	NdeI
218	Forward	CGCGGATCCCATATG-GTCGCGGTCGATC	XhoI
	Reverse	CCCGCTCGAG-TAACTCATAGAATCCTGC	BamHI-
219	Forward	CGCGGATCCGCTAGC-ACGGCAAGGTTAAG	NdeI
	Reverse	CCCGCTCGAG-TTTAAACCATCTCCTCAAAAC	XhoI
223	Forward	CGCGGATCCCATATG-GAATTCAGGCACCAAGTA	BamHI-
	Reverse	CCCGCTCGAG-GGCTTCCCGCGTGTC	NdeI
225	Forward	CGCGGATCCCATATG-GACGAGTTGACCAACC	XhoI
	Reverse	CCCGCTCGAG-GTTCAGAAAGCGGGAC	BamHI-
226	Forward	AAAGAATTC-CTTGCGATTATCGTGCGCACGCG	NdeI
	Reverse	AAACTGCAG-TCAAAATCCCAAAACGGGGAT	XhoI
228	Forward	CGCGGATCCCATATG-TCGCAAGAAGCCAAACAG	Eco RI
	Reverse	CCCGCTCGAG-TTTGGCGGCATCTTTCAT	Pst I
229	Forward	CGCGGATCCCATATG-CAAGAGGTTTGGCC	BamHI-
	Reverse	CCCGCTCGAG-ACACAATATAGCGGATGAAC	NdeI
230	Forward	CGCGGATCCCATATG-CATCCGGGTGCCGAC	XhoI
	Reverse	CCCGCTCGAG-AAGTTTGGCGGCTTCGG	BamHI-
232	Forward	AAAAAAGAATTC-ATGTACGCTAAAAAAGGCGGTTTGGG	NdeI
	Reverse	AAAAAACTGCAG-TCAAGGTTTTTCTGATTGCCGCCGC	XhoI
232a	Forward	AAAAAAGAATTC-GCCAAGGCTGCCGATACACAAATTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACATTGTCGTTGCCGCCAGATG	Pst I
233	Forward	CGCGGATCCCATATG-GCGGACAAACCCAAG	BamHI-
	Reverse	CCCGCTCGAG-GACGGCATTGAGCAG	NdeI
234	Forward	CGCGGATCCCATATG-GCCGTTTCACTGACCG	XhoI
			BamHI-

	Reverse	GCCCAAGCTT-ACGGTTGGATTGCCATG	NdeI
235	Forward	CGCGGATCCCATATG-GCCTGCCAAGTTCAAA	Hind III
	Reverse	CCCGCTCGAG-TTTGGGCTGCTCTTC	BamHI-
236	Forward	CGCGGATCCCATATG-GCGCGTTTCGCCTT	NdeI
	Reverse	CCCGCTCGAG-ATGGGTCGCGCGCCGT	XhoI
238	Forward	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCCG	BamHI-
	Reverse	CCCGCTCGAG-TTTGTCTAAGTTCCTGATATG	NdeI
239	Forward	CCGGAATTCTACATATG-CTCCACCATAAAGGTATTG	XhoI
	Reverse	CCCGCTCGAG-TGGTGAAGAGCGGTTTAG	EcoRI-
240	Forward	CGCGGATCCCATATG-GACGTTGGACGATTTC	NdeI
	Reverse	CCCGCTCGAG-AAACGCCATTACCCGATG	XhoI
241	Forward	CCGGAATTCTACATATG-CCAACACGTCCAAC	BamHI-
	Reverse	CCCGCTCGAG-GAATGCGCCTGTAATTAATC	NdeI
242	Forward	CGCGGATCCCATATG-ATCGGCAAACCTTGTTG	XhoI
	Reverse	GCCCAAGCTT-ACCGATACGGTCGCAG	BamHI-
243	Forward	CGCGGATCCCATATG-ACGATTTTTTCGATGCTGC	NdeI
	Reverse	CCCGCTCGAG-CGACTTGGTTACCGCG	HindIII
244	Forward	CGCGGATCCCATATG-CCGTCTGAAGCCC	BamHI-
	Reverse	CCCGCTCGAG-TTTTTTCGGTAGGGGATTT	NdeI
246	Forward	CGCGGATCCCATATG-GACATCGGCAGTGC	XhoI
	Reverse	CCCGCTCGAG-CCCGCGCTGCTGGAG	BamHI-
247	Forward	CGCGGATCCCATATG-GTCGGATCGAGTTAC	NdeI
	Reverse	CCCGCTCGAG-AAGTGTTCTGTTTGCGCA	XhoI
248	Forward	CGCGGATCCCATATG-CGCAAACAGAACT	BamHI-
	Reverse	CCCGCTCGAG-CTCATCATTATTGCTAACA	NdeI
249	Forward	CGCGGATCCCATATG-AAGAATAATGATTGCTTC	XhoI
	Reverse	CCCGCTCGAG-TTCCCGACCTCCGAC	BamHI-
251	Forward	CGCGGATCCCATATG-CGTGCTGCGGTAGT	NdeI
	Reverse	CCCGCTCGAG-TACGAAAGCCGGTCGTG	XhoI
253	Forward	AAAAAAGAATTC-ATGATTGACAGGAACCGTATGCTGCG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I

253a	Forward	AAAAAAGAATTC-AAAATCCTTTTGAAAACAAGCGAAAACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTGGTCTTTCAAACGCCCTTCCTG	Pst I
254	Forward	AAAAAAGAATTC-ATGTATACAGGCGAACGCTTCAATAC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATTACGTAACCGTACACGCTGAC	Xba I
255	Forward	CGCGGATCCCATATG-GCCGCGTTGCGTTAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCCGCAATACCGACCAG	XhoI
256	Forward	CGCGGATCCGCTAGC-TTTTAACACCGCCGGAC	BamHI- NheI
	Reverse	CCCGCTCGAG-ACGCCTGTTTGTGCGG	XhoI
257	Forward	CGCGGATCCCATATG-GCGGTTTCTTTCCTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCGTGAATATCGCG	XhoI
258	Forward	AAAAAAGAATTC-GATTATTTCTGGTGGATTGTTGCGTTCAG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
258a	Forward	AAAAAAGAATTC-GCGAAGGCGGTGGCGCAAGGCGA	Eco RI
	Reverse	AAAAAACTGCAG-CTACGCATAAGTTTTTACCGTTTTTGG	Pst I
259	Forward	CGCGGATCCCATATG-GAAGAGCTGCCTCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCTTTTCCGGCGTTT	XhoI
260	Forward	CGCGGATCCCATATG-GGTGCGGGTATGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACAGGGCGACACCCT	XhoI
261	Forward	AAAAAAGAATTC-CAAGATACAGCTCGGGCATTTCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAAACCAACAAGCCTTGGTCACT	Pst I
263	Forward	CGCGGATCCCATATG-GCACGTTTAACCGTA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGTAAGCCTGCAATT	XhoI
264	Forward	AAAAAAGGTACC-GCCGACGCAGTGGTCAAGGCAGAA	Kpn I
	Reverse	AAACTGCAG-TCAGCCGGCGGTCAATACCGCCCG	Pst I
265	Forward	AAAAAAGAATTC-GCGGAGGTCAAGAGAAGGTGTTTG	Eco RI
	Reverse	AAAAAACTGCAG-TTACGAATACGTCGTCAAAATGGG	Pst I
266	Forward	AAAGAATTC-CTCATCTTTGCCAACGCCCCCTTC	Eco RI
	Reverse	AAACTGCAG-CTATTCCCTGTTGCGCGTGTGCCA	Pst I
267	Forward	AAAGAATTC-TTCTTCCGATTTCGATGTTAATCG	Eco RI
	Reverse	AAACTGCAG-TTAGTAAAAACCTTTCTGCTTGGC	Pst I
269	Forward	AAAGAATTC-TGCAAACCTTGCGCCACGTGCCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
269a	Forward	AAAAAAGAATTC-GACTTTATCCAAAACACGGCTTCGCC	Eco RI
	Reverse	AAACTGCAG-TTACGAAGACCGCAACGAAAGGCAGAG	Pst I
270	Forward	AAAGAATTC-GCCGTCAAGCTCGTTTTGTGCAATG	Eco RI
	Reverse	AAACTGCAG-TTATTCGGCGGTAAATGCCGTCTG	Pst I
271	Forward	CGCGGATCCCATATG-CCTGTGTGCAGCTCGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCCCAGCCCCGTGGAG	XhoI

272	Forward	AAAGAATTC-ATGACCGCAAAGGAAGAACTGTTTCGC	Eco RI
	Reverse	AAACTGCAG-TCAGAGCAGTTCCAAATCGGGGCT	Pst I
273	Forward	AAAGAATTC-ATGAGTCTTCAGGCGGTATTTATATACCC	Eco RI
	Reverse	AAACTGCAG-TTACGCGTAAGAAAAAACTGC	Pst I
274	Forward	CGCGGATCCCATATG-ACAGATTGGTTACGGAC	BamHI-
	Reverse	CCCGCTCGAG-TTTGCTTTTCA GTATTATTGAA	NdeI
			XhoI
276	Forward	AAAAAAGAATTC- ATGATTTTGCCGTCGTCCATCACGATGATGCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACACCACCATCGGCGAATTTATGGC	Pst I
277	Forward	AAAAAAGAATTC-ATGCCCCGCTTTGAGGACAAGCTCGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
277a	Forward	AAAAAAGAATTC-GGGGCGGCGGCTGGGTTGGACGTAGG	Eco RI
	Reverse	AAAAAACTGCAG-TCATAAGCCATGCTTACCTTCCAACAA	Pst I
278	Forward	AAAAAAGGTACC-GTCAAAGTTGTATTAATCGGGCCTTTGCC	Kpn I
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
278a	Forward	AAAAAAGAATTC-AAAACTCTCCTAATTCGTCATAGTCG	Eco RI
	Reverse	AAAAAACTGCAG-TCATTCAACCATATCAAATCTGCC	Pst I
279	Forward	CGCGGATCCCATATG-TTGCTGCAATCACGATT	BamHI-
	Reverse	CCCGCTCGAG-TTTAGAAGCGGGCGGCAA	NdeI
			XhoI
280	Forward	AAAAAAGGTACC-GCCCCCTGCCGTTGTAACCG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTGCTTCATCGCGTTGGTCAAGGC	Pst I
281	Forward	AAAAAAGAATTC-GCACCCGTCGGCGTATTCTCGTCATGCG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
281a	Forward	AAAAAAGAATTC-TCCTACCACATCGAAATTCCTTCCGG	Eco RI
	Reverse	AAAAAATCTAGA-GGTCAGAATGCCGCCTTCTTTGCCGAG	Xba I
282	Forward	AAAAAAGAATTC-CTTTACCTTGACCTGACCAACGGGCACAG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACCTGCCAGTTGCGGGAATATCGT	Pst I
283	Forward	CGCGGATCCCATATG-GCCGTCTTTACTTGGAAG	BamHI-
	Reverse	CCCGCTCGAG-ACGGCAGTATTTGTTTACG	NdeI
			XhoI
284	Forward	CGCGGATCCCATATG-TTGCCTGCAAAAGAATCG	BamHI-
	Reverse	CCCGCTCGAG-CCGACTTTGCAAAAAGT	NdeI
			XhoI
286	Forward	CGCGGATCCCATATG-GCCGACCTTTCCGAAAA	BamHI-
	Reverse	CCCGCTCGAG-GAAGCGCGTTCCCAAG	NdeI
			XhoI
287	Forward	CCGGAATTCTAGCTAGC-CTTTCAGCCTGCGGG	EcoRI-
	Reverse	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC	NheI
			XhoI
288	Forward	CGCGGATCCCATATG-CACACCGGACAGG	BamHI-
	Reverse	CCCGCTCGAG-CGTATCAAAGACTTGCGT	NdeI
			XhoI
290	Forward	CGCGGATCCCATATG-GCGGTTTGGGGCGGA	BamHI-

	Reverse	CCCGCTCGAG-TCGGCGCGGCGGGC	NdeI
292	Forward	CGCGGATCCCATATG-TGCGGGCAAACGCCC	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-TTGATTTTTGCGGATGATTT	NdeI
294	Forward	AAAAAAGAATTC-GTCTGGTCGATTCGGGTTGTCAGAAC	XhoI
	Reverse	AAAAAACTGCAG-TTACCAGCTGATATAAAACATCGCTTT	Eco RI
295	Forward	CGCGGATCCCATATG-AACCGGCCGGCCTCC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CGATATTTGATTCCGTTGC	NdeI
297	Forward	AAAAAAGAATTC-GCATACATTGCTTCGACAGAGAG	XhoI
	Reverse	AAAAAACTGCAG-TCAATCCGATTGCGACACGGT	Eco RI
298	Forward	AAAAAAGAATTC-CTGATTGCCGTGTGGTTCAGCCAAAACCC	Pst I
	Reverse	AAAAAACTGCAG-TCATGGCTGTGTACTTGATGGTTGCGT	Eco RI
299	Forward	CGCGGATCCGCTAGC-CTACCTGTCGCCTCCG	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-TTGCCTGATTGCAGCGG	NheI
302	Forward	AAAAAAGAATTC-ATGAGTCAAACCGATACGCAACG	XhoI
	Reverse	AAAAAACTGCAG-TTAAGGTGCGGGATAGAATGTGGGCGC	Eco RI
305	Forward	AAAAAAGGTACC-GAATTTTACCGATTTCCAGCACCGGA	Pst I
	Reverse	AAAAAACTGCAG-TCATTCCTCACTTATCCAGCCTGACAG	Kpn I
305a	Forward	AAAAAAGGTACC-TCCCGTTCGGGCAGTACGATTATGGG	Pst I
	Reverse	AAAAAACTGCAG-TTACAAACCGACATCATGCAGGGTGAA	Kpn I
306	Forward	CGCGGATCCCATATG-TTTATGAACAAATTTTCCC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CCGCATCGGCAGAC	NdeI
308	Forward	CGCGGATCCCATATG-TTAAATCGGGTATTTTATC	XhoI
			BamHI-
	Reverse	CCCGCTCGAG-ATCCGCCATTCCCTGC	NdeI
311	Forward	AAAAAAGGTACC-ATGTTTCAGTTTTGGCTGGGTGTTT	XhoI
	Reverse	AAACTGCAG-ATGTTTCATATTCCTGCCTTCGGC	Kpn I
312	Forward	AAAAAAGGTACC-ATGAGTATCCCATCCGGCGAAATT	Pst I
	Reverse	AAACTGCAG-TCAGTTTTTCATCGATTGAACCGG	Kpn I
313	Forward	AAAAAAGAATTC-ATGGACGACCCGCGCACCTACGGATC	Pst I
	Reverse	AAAAAACTGCAG-TCAGCGGCTGCCGCCGATTTTGCT	Eco RI
401	Forward	CGCGGATCCCATATG-AAGGCGGCAACACAGC	Pst I
			BamHI-
	Reverse	CCCGCTCGAG-CCTTACGTTTTTCAAAGCC	NdeI
402	Forward	AAAAAAGAATTC-GTGCCTCAGGCATTTTCATTTACCCTTGC	XhoI
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Eco RI
402a	Forward	AAAAAAGAATTC-AGGCTGATTGAAAACAAACACGG	Xba I
	Reverse	AAAAAATCTAGA-TTAAATCCCTCTGCCGTATTTGTATTC	Eco RI
406	Forward	CGCGGATCCCATATG-TGCGGGACACTGACAG	Xba I
			BamHI-
	Reverse	CCCGCTCGAG-AGGTTGTCCTTGTCTATG	NdeI
			XhoI

501	Forward	CGCGGATCCCATATG-GCAGGCGGAGATGGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGTGTGATGTTACCC	XhoI
502	Forward	CGCGGATCCCATATG-GTAGACGCGCTTAAGCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCTGCATGGCGGCG	XhoI
503	Forward	CGCGGATCCCATATG-TGTTCTGGGGAAAGGCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CCGCGCATTCTCGCA	XhoI
504	Forward	CGCGGATCCCATATG-AGCGATATTGAAGTGACG	BamHI- NdeI
	Reverse	GCCCAAGCTT-TGATTCAAGTCCTTGCCG	HindIII
505	Forward	CGCGGATCCCATATG-TTTCGTTTACAATTCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGGCGTTTTATAGCGG	XhoI
510	Forward	CGCGGATCCCATATG-CCTTCGCGGACAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCACTGGCAGCG	XhoI
512	Forward	CGCGGATCCCATATG-GGACATGAAGTAACGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGAATAGCCTTTGACG	XhoI
515	Forward	CGCGGATCCCATATG-GAGGAAATAGCCTTCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATGCCGCAAAGCATC	XhoI
516	Forward	CGCGGATCCCATATG-TGTACGTTGATGTTGTGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGCGGGCGGCATC	XhoI
517	Forward	CGCGGATCCCATATG-GGTAAAGGTGTGGAAATA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGCGCCCAGCCGT	XhoI
518	Forward	AAAGAATTC-GCTTTTTTACTGCTCCGACCGGAAGG	Eco RI
	Reverse	AAACTGCAG-TCAAATTCAGACTCTGCCAC	Pst I
519	Forward	CGCGGATCCCATATG-TTCAAATCCTTTGTCGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGGCGGTTTTGCTGC	XhoI
520	Forward	CGCGGATCCCATATG-CCTGCGCTTCTTTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATATTTACATTTAGTCGGC	XhoI
521	Forward	CGCGGATCCCATATG-GCCAAAATCTATACCTGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATACGCCCCAGTTCC	XhoI
522	Forward	CGCGGATCCCATATG-ACTGAGCCGAAACAC	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTCTGATTTCAAATCGGCA	HindIII
523	Forward	CGCGGATCCCATATG-GCTCTGCTTTCCGCG	BamHI- NdeI

	Reverse	CCCGCTCGAG-AGGGTGTGTGATAATAAGAAG	XhoI
525	Forward	CGCGGATCCCATATG-GCCGAAATGGTTCAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCCGTGCATATCATAAA	XhoI
527	Forward	AAAGAATTC-TTCCCTCAATGTTGCCGTTTTTCG	Eco RI
	Reverse	AAACTGCAG-TTATGCTAAACTCGAAACAAATTC	Pst I
529	Forward	CGCGGATCCGCTAGC-TGCTCCGGCAGCAAAAC	BamHI- NheI
	Reverse	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
530	Forward	CGCGGATCCCATATG-AGTGCGAGCGCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGACCGACTGATTCCG	XhoI
531	Forward	AAAAAAGAATTC-TATGCCGCCGCTACCAAATCTACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAAACAGCGCCGTGCCGACGACAAG	Pst I
532	Forward	AAAAAAGAATTC-ATGAGCGGTCAGTTGGGCAAAGGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCAGTGGTCGGTATCAAA	Pst I
532a	Forward	AAAAAAGAATTC-TTGGGTGTCGCGTTTGAGCCGGAAGT	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTGTTCAGTGGTCGGTATCAAA	Pst I
535	Forward	AAAGAATTC-ATGCCCTTTCCCGTTTTTCAGAC	Eco RI
	Reverse	AAACTGCAG-TCAGACGACCCCGCCTTCCCC	Pst I
537	Forward	CGCGGATCCCATATG-CATACCCAAAACCAATCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATCCTGCAAATAAAGGGTT	XhoI
538	Forward	CGCGGATCCCATATG-GTCGAGCTGGTCAAAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGGCATTTCGGTTTCGTC	XhoI
539	Forward	CGCGGATCCGCTAGC-GAGGATTTGCAGGAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TACCAATGTCGGCAAATC	XhoI
542	Forward	AAAGAATTC-ATGCCGTCTGAAACCGTGTC	Eco RI
	Reverse	AAACTGCAG-TTACCGCGAACC GGTCAGGAT	Pst I
543	Forward	AAAAAAGAATTC-GCCTTCGATGGCGACGTTGTAGGTAC	Eco RI
	Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
543a	Forward	AAAAAAGAATTC-GGCAAAACTCGTCATGAATTTGC	Eco RI
	Reverse	AAAAAATCTAGA- TTAATGAAGAAGAACATATTGGAATTTTGG	Xba I
544	Forward	AAAGAATTC-GCGCCCGCCTTCTCCCTGCCCCGACCTGCACGG	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
544a	Forward	AAAAAAGAATTC- GCAAATGACTATAAAAAACAAAACTTCCAAGTACTTGC	Eco RI
	Reverse	AAACTGCAG-CTATTGCGCCACGCGCGTATCGAT	Pst I
547	Forward	AAAGAATTC-ATGTTCTGTAGATAACGGATTTAATAAAAC	Eco RI
	Reverse	AAACTGCAG-TTAACAACAAAAACAAACCGCTT	Pst I
548	Forward	AAAGAATTC-GCCTGCAAACCTCAAGACAACAGTGCGGC	Eco RI

Reverse	AAACTGCAG-TCAGAGCAGGGTCCTTACATCGGC	Pst I
550 Forward	AAAAAAGTTCGAC- ATGATAACGGACAGGTTTCATCTCTTTTCATTTTCC	Sal I
Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
550a Forward	AAAAAAGAATTC-GTAAATCACGCCTTTGGAGTCGCAAACGG	Eco RI
Reverse	AAACTGCAG-TTACGCAAACGCTGCAAAATCCCC	Pst I
552 Forward	AAAAAAGAATTC-TTGGCGCGTTGGCTGGATAC	Eco RI
Reverse	AAACTGCAG-TTATTTCTGATGCCTTTTCCCAAC	Pst I
554 Forward	CGCGGATCCCATATG-TCGCCCCGCGCCAAC	BamHI- NdeI
Reverse	CCCGCTCGAG-CTGCCCTGTCAGACAC	XhoI
556 Forward	AAAGAATTC-GCGGGCGGTTTGTGTTGGACATCCCCG	Eco RI
Reverse	AAACTGCAG-TTAACGGTGCGGACGTTTCTGACC	Pst I
557 Forward	CGCGGATCCCATATG-TGCGGTTTCCACCTGAA	BamHI- NdeI
Reverse	CCCGCTCGAG-TTCCGCCTTCAGAAAGG	XhoI
558 Forward	AAAGAATTC-GAGCTTTATATGTTTCAACAGGGGACGGC	Eco RI
Reverse	AAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
558a Forward	AAAAAAGAATTC-ATTAGATTCTATCGCCATAAACAGACGGG	Eco RI
Reverse	AAAAAACTGCAG-CTAAACAATGCCGTCTGAAAGTGGAGA	Pst I
560 Forward	AAAAAAGAATTC- TCGCCTTTCCGGGACGGGGCGCACAAGATGGC	Eco RI
Reverse	AAAAAACTGCAG-TCATGCGGTTTTCAGACGGCATTGTTGGC	Pst I
561 Forward	CCGGAATTCTACATATG-ATACTGCCAGCCCCGT	EcoRI- NdeI
Reverse	CCCGCTCGAG-TTTCAAGCTTTCTTCAGATG	XhoI
562 Forward	CGCGGATCCCATATG-GCAAGCCCGTCGAG	BamHI- NdeI
Reverse	CCCGCTCGAG-AGACCAACTCCAACCTCGT	XhoI
565 Forward	CGCGGATCCCATATG-AAGTCGAGCGCGAAATAC	BamHI- NdeI
Reverse	CCCGCTCGAG-GGCATTGATCGGCGGC	XhoI
566 Forward	CGCGGATCCCATATG-GTCGGTGGCGAAGAGG	BamHI- NdeI
Reverse	CCCGCTCGAG-CGCATGGGCGAAGTCA	XhoI
567 Forward	CCGGAATTCTACATATG-AGTGCGAACATCCTTG	EcoRI- NdeI
Reverse	CCCGCTCGAG-TTTCCCCGACACCCTCG	XhoI
568 Forward	CGCGGATCCCATATG-CTCAGGGTCAGACC	BamHI- NdeI
Reverse	CCCGCTCGAG-CGGCGCGGCGTTTCAG	XhoI
569 Forward	AAAAAAGAATTC-CTGATTGCCTTGTGGGAATATGCCCG	Eco RI
Reverse	AAAAAACTGCAG-TTATGCATAGACGCTGATAACGGCAAT	Pst I
570 Forward	CGCGGATCCCATATG-GACACCTTCCAAAAAATCG	BamHI- NdeI
Reverse	CCCGCTCGAG-GCGGGCGTTTCATTTCTTT	XhoI

571	Forward	AAAAAAGAATTC- ATGGGTATTGCCGGCGCCGTAAATGTTTTGAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TTATGGCCGACGCGCGGCTACCTGACG	Pst I
572	Forward	CGCGGATCCCATATG-GCGCAAAAAGGCAAAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGCAGTGTGCCGATA	XhoI
573	Forward	CGCGGATCCCATATG-CCCTGTTTGTGCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GACGGTGTTCATTTCCGC	XhoI
574	Forward	CGCGGATCCCATATG-TGGTTTGCCGCCCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACTTCGATTTTATTCGGG	XhoI
575	Forward	CGCGGATCCCATATG-GTTTCGGGCGAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCCGAATCTGAACAG	XhoI
576	Forward	CGCGGATCCCATATG-GCCGCCCCCGCATCT	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTACTTTTTTGATGTGCAC	XhoI
577	Forward	CGCGGATCCCATATG-GAAAGGAACGGTGTATTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGGCTGTTTGGTAGATTCTG	XhoI
578	Forward	CGCGGATCCCATATG-AGAAGGTTTCGTACAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCAACGCCTCCACG	XhoI
579	Forward	CGCGGATCCCATATG-AGATTGGGCGTTTCCAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGAATTGATGATGTGTATGT	XhoI
580	Forward	CGCGGATCCCATATG-AGGCAGACTTCGCCGA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CACTTCCCCCGAAGTG	XhoI
581	Forward	CGCGGATCCCATATG-CACTTCGCCCAGC	BamHI- NdeI
	Reverse	CCCGCTCGAG-CGCCGTTTGGCTTTGG	XhoI
582	Forward	AAAAAAGAATTC-TTTGGAGAGACCGCGCTGCAATGCGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGATGCCGTCCCAGTCGTTGAA	Xba I
583	Forward	AAAAAAGAATTC-ACTGCCGGCAATCGACTGCATAATCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGAGGTCAATATGATGAAATTG	Pst I
584	Forward	AAAAAAGAATTC- GCGGCTGAAGCATTGAATTACAATATTGTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGAACTGAACCGTCCCATTGACGCT	Pst I
585	Forward	AAAAAAGGTACC-TCTTTCTGGCTGGTGCAGAACACCCTTGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGTTCGCACTTTTTTCTGTTTTGGA	Pst I
586	Forward	CGCGGATCCCATATG-GCAGCCCATCTCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTCAGCGAATCAAGTTTC	XhoI
587	Forward	CGCGGATCCCATATG-GACCTGCCCTTGACGA	BamHI-

	Reverse	CCCGCTCGAG-AAATGTATGCTGTACGCC	NdeI
			XhoI
588	Forward	AAAAAAGAATTC-GCCGTCCTGACTTCCTATCAAGAACCAGG	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTGTTTTTGGGCAGTTTCACTTC	Pst I
589	Forward	AAAAAAGAATTC- ATGCAACAAAAAATCCGTTTCCAAATCGAAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAATCGATTTTACCCGTTTCAGGCG	Pst I
590	Forward	AAAAAAGAATTC-ATGAAAAAACCTTTGATFTCAGTTGCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACTGCTGCGGCTCTGAAACCAT	Pst I
591	Forward	AAAAAAGAATTC-CACTACATCGTTGCCAGATTGTGCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTAACCGAGCAGCCGGGTAACGTCGTT	Pst I
592a	Forward	AAAAAAGAATTC-CGCGATTACACCGCCAAGCTGAAAATGGG	Eco RI
	Reverse	AAAAAACTGCAG-TTACCAAACGTCGGATTGTGATACG	Pst I
593	Forward	CGCGGATCCGCTAGC-CTTGAACCTGAACGGACTC	BamHI-
			NheI
	Reverse	CCCGCTCGAG-GCGGAAGCGGACGATT	XhoI
594a	Forward	AAAAAAGAATTC-GGTAAGTTCGCCGTTTCAGGCCTTTCA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCCGCCGTTTCTGACACTCGCG	Pst I
595	Forward	AAAAAAGAATTC-TGCCAGCCGCCGAGGCGGAGAAAGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCAAGCCGAGTATGCCGCG	Pst I
596	Forward	CGCGGATCCCATATG-TCCCAACAATACGTC	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-ACGCGTTACCGGTTTGT	XhoI
597	Forward	CGCGGATCCCATATG-CTGCTTCATGTCAGC	BamHI-
			NdeI
	Reverse	GCCCAAGCTT-ACGTATCCAGCTCGAAG	HindIII
601	Forward	CGCGGATCCCATATG-ATATGTTCCCAACCGGCAAT	BamHI-
			NdeI
	Reverse	CCCGCTCGAG-AAAACAATCCTCAGGCAC	XhoI
602	Forward	CGCGGATCCGCTAGC-TTGCTCCATCAATGC	BamHI-
			NheI
	Reverse	CCCGCTCGAG-ATGCAGCTGCTAAAAGCG	XhoI
603	Forward	AAAAAAGAATTC-CTGTCCTCGCGTAGGCGGGGACGGGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAGATGCCGGCAAGTTCGGC	Pst I
604	Forward	CGCGGATCCGCTAGC-CCCGAAGCGCACTT	BamHI-
			NheI
	Reverse	CCCGCTCGAG-GACGGCATCTGCACGG	XhoI
606a	Forward	AAAAAAGAATTC-CGCGAATACCGCGCCGATGCGGGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTAAAGCGATTTGAGGCGGGCGATACG	Pst I
607	Forward	AAAAAAGAATTC-ATGCTGCTCGACCTCAACCGCTTTTC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCCTTATGCGATCTGAC	Pst I
608	Forward	AAAAAAGAATTC-ATGTCCGCCCTCCTCCCCATCATCAACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGTCTATCCAAATGTCGCGTTC	Pst I
609	Forward	CGCGGATCCCATATG-GTTGTGGATAGACTCG	BamHI-
			NdeI

	Reverse CCCGCTCGAG-CTGGATTATGATGTCTGTC	XhoI
610	Forward CGCGGATCCCATATG-ATTGGAGGGCTTATGCA	BamHI- NdeI
	Reverse CCCGCTCGAG-ACGCTTCAACATCTTTGCC	XhoI
611	Forward CGCGGATCCCATATG-CCGTCTCAAAACGGG	BamHI- NdeI
	Reverse CCCGCTCGAG-AACGACTTTGAACGCGCAA	XhoI
613	Forward CGCGGATCCCATATG-TCGCGTTCGAGCCG3-	BamHI- NdeI
	Reverse CCCGCTCGAG-AGCCTGTAAAATAAGCGGC	XhoI
614	Forward CGCGGATCCCATATG-TCCGTCGTGAGCGGC	BamHI- NdeI
	Reverse CCCGCTCGAG-CCATACTGCGGCGTTC	XhoI
616	Forward AAAAAAGAATTC-ATGTCAAACACAATCAAAATGGTTGTGCGG	Eco RI
	Reverse AAAAAATCTAGA-TTAGTCCGGGCGGCAGGCAGCTCG	Xba I
619a	Forward AAAAAAGAATTC-GGGCTTCTCGCCGCTCGCTTG	Eco RI
	Reverse AAAAAACTGCAG-TCATTTTTTGTGTTTTAAACGAGATA	Pst I
622	Forward CGCGGATCCCATATG-GCCGCCCTGCCTAAAG	BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGTCCAAATGATAAATCTG	XhoI
624	Forward CGCGGATCCCATATG-TCCCCGCGCTTTTACCG	BamHI- NdeI
	Reverse CCCGCTCGAG-AGATTCCGGGCTGCGC	XhoI
625	Forward CGCGGATCCCATATG-TTTGCAACCAGGAAAATG	BamHI- NdeI
	Reverse CCCGCTCGAG-CGGCAAAATTACCGCCTT	XhoI
627a	Forward AAAAAAGAATTC-AAAGCAGGCGAGGCAGGCGCGCTGGG	Eco RI
	Reverse AAAAAACTGCAG- TTACGAATGAAACAGGGTACCCGTCATCAAGGC	Pst I
628	Forward AAAAAAGGTACC-GCCTTACAAACATGGATTTTGCGTTC	Kpn I
	Reverse AAAAAACTGCAG-CTACGCACCTGAAGCGCTGGCAA	Pst I
629a	Forward AAAAAAGAATTC-GCCACCTTTATCGCGTATGAAAACGA	Eco RI
	Reverse AAAAAACTGCAG-TTACAACACCGCCGTCCGGTTCAAACC	Pst I
630a	Forward AAAAAAGAATTC-GCGGCTTTGGGTATTTCTTTCGG	Eco RI
	Reverse AAAAAACTGCAG-TTAGGAGACTTCGCCAATGGAGCCGGG	Pst I
635	Forward AAAAAAGAATTC- ATGACCCAGCGACGGGTGCGCAAGCAAAACCG	Eco RI
	Reverse AAAAAACTGCAG-TTAATCCACTATAATCCTGTTGCT	Pst I
638	Forward AAAAAAGAATTC-ATGATTGGCGAAAAGTTTATCGTAGTTGG	Eco RI
	Reverse AAAAAACTGCAG-TCACGAACCGATTATGCTGATCGG	Pst I
639	Forward CGCGGATCCCATATG-ATGCTTTATTTGTTCG	BamHI- NdeI
	Reverse CCCGCTCGAG-ATCGCGGCTGCCGAC	XhoI
642	Forward CGCGGATCCCATATG-CGGTATCCGCCGCAAT	BamHI- NdeI
	Reverse CCCGCTCGAG-AGGATTGCGGGGCATTA	XhoI

643	Forward	CGCGGATCCCATATG-GCTTCGCCGTCGGCAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACCGAAAAACAGACCGC	XhoI
644	Forward	AAAAAAGAATTC- ATGCCGTCTGAAAGGTCGGCGGATTGTTGCCC	Eco RI
	Reverse	AAAAAATCTAGA-CTACCCGCAATATCGGCAGTCCAATAT	Pst I
645	Forward	AAAAAAGAATTC-GTGAACAGAGCAACACGTAAATCG	Eco RI
	Reverse	AAAAAACTGCAG-CTACGAGGAAACCGAAGACCAGGCCGC	Pst I
647	Forward	AAAAAAGAATTC-ATGCAAAGGCTCGCCGCAGACGG	Eco RI
	Reverse	AAAAAACTGCAG-TTAGATTATCAGGGATATCCGGTAGAA	Pst I
648	Forward	AAAAAAGAATTC- ATGAACAGGCGCGACGCGCGGATCGAACG	Eco RI
	Reverse	AAAAAACTGCAG-TCAAGCTGTGTGCTGATTGAATGCGAC	Pst I
649	Forward	AAAAAAGAATTC-GGTACGTCAGAACCCGCCACCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGGCGGAAACTGCCGCCGTC	Pst I
650	Forward	AAAAAAGAATTC-ATGTCCAAACTCAAACCATCGC	Eco RI
	Reverse	AAAAAACTGCAG-TCAGACGGCATGGCGGTCTGTTTT	Pst I
652	Forward	AAAAAAGGTACC- GCTGCCGAAGACTCAGGCCTGCCGCTTTACCG	Kpn I
	Reverse	AAAAAACTGCAG-TTATTTGCCAGTTGGTAGAATGCGGC	Pst I
653	Forward	AAAAAAGAATTC-GCGGCTTTGCCGGTAATTTTCATCGG	Eco RI
	Reverse	AAAAAACTGCAG-CTATGCCGGTCTGGTTGCCGGCGGCGA	Pst I
656a	Forward	AAAAAAGAATTC-CGGCCGACGTCGTTGCGTCCTAAGTC	Eco RI
	Reverse	AAAAAACTGCAG-CTACGATTCGGCGATTTCACATCGT	Pst I
657	Forward	AAAAAAGAATTC-GCAGAATTTGCCGACCGCCATTTGTGCGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATAGGGACTGATGCAGTTTTTTTGC	Pst I
658	Forward	CGCGGATCCCATATG-GTGTCCGGAATTGTG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCAGAATGTTTACCGTT	XhoI
661	Forward	AAAAAAGAATTC- ATGCACATCGGCGGCTATTTTATCGACAACCC	Eco RI
	Reverse	AAAAAACTGCAG-TCACGACGTGTCTGTTCCGCCGTCGGGC	Pst I
663	Forward	CGCGGATCCCATATG-TGTATCGAGATGAAATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTAAAAATCGGGGCTGC	XhoI
664	Forward	CGCGGATCCCATATG-GCGGCTGGCGCGGT	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAATCGAGTTTTACACCAC	XhoI
665	Forward	AAAAAAGAATTC-ATGAAATGGGACGAAACGCGCTTCGG	Eco RI
	Reverse	AAAAAACTGCAG-TCAATCCAAAATTTTGCCGACGATTTC	Pst I
666	Forward	AAAAAAGAATTC-AACTCAGGCGAAGGAGTGCTTGTGGC	Eco RI
	Reverse	AAAAAATCTAGA-TCAGTTTAGGGATAGCAGGCGTAC	Xba I
667	Forward	AAAAAAGAATTC- CCGCATCCGTTTGATTTCATTTCGTATTCGTCCG	Eco RI
	Reverse	AAAAAACTGCAG-TTAATGACACAATAGGCGCAAGTC	Pst I

669	Forward	AAAAAAGAATTC-ATGCGCCGCATCATTAACACACCAGCC	Eco RI
	Reverse	AAAAAACTGCAG-TTACAGTATCCGTTTGTATGTCGGC	Pst I
670a	Forward	AAAAAAGAATTC-AAAAACGCTTCGGGCGTTTCGTCTTC	Eco RI
	Reverse	AAAAAACTGCAG- TTAGGAGCTTTTGGAACGCGTCGGACTGGC	Pst I
671	Forward	CGCGGATCCCATATG-ACCAGCAGGGTAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AGCAACTATAAAAAACGCAAG	XhoI
672	Forward	CGCGGATCCCATATG-AGGAAAATCCGCACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-ACGGGATAGGCGGTTG	XhoI
673	Forward	AAAAAAGAATTC-ATGGATATTGAAACCTTCCTTGCAGG	Eco RI
	Reverse	AAAAAACTGCAG-CTACAAACCCAGCTCGCGCAGGAA	Pst I
674	Forward	AAAAAAGAATTC-ATGAAAACAGCCCGCCCGCTTCCCG	Eco RI
	Reverse	AAAAAACTGCAG-TCAACGGCGTTTGGGCTCGTCGGG	Pst I
675	Forward	CGCGGATCCCATATG-AACACCATCGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTCTTCGTCTTCAAACCTGT	XhoI
677a	Forward	AAAAAAGAATTC-AGACGGCATTCCCGATCAGTCGATTTTGA	Eco RI
	Reverse	AAAAAACTGCAG-TTACGTATGCGCGAAATCGACCGCCGC	Pst I
680	Forward	CGCGGATCCGCTAGC-ACGAAGGGCAGTTCGG	BamHI- NheI
	Reverse	CCCGCTCGAG-CATCAAAAACCTGCCGC	XhoI
681	Forward	AAAAAAGAATTC-ATGACGACGCCGATGGCAATCAGTGC	Eco RI
	Reverse	AAAAAACTGCAG-TTACCGTCTTCCGCAAAAACAGC	Pst I
683	Forward	CGCGGATCCCATATG-TGCAGCACACCGGACAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAGTTTTTTTCCGCATACG	XhoI
684	Forward	CGCGGATCCCATATG-TGCGGTACTGTGCAAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTCGACCATCTGTTGCG	XhoI
685	Forward	CGCGGATCCCATATG-TGTTTGCTTAATAATAAACATT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTTTTTTCCCGCCGCA	XhoI
686	Forward	CGCGGATCCCATATG-TGCGGCGGTTTCGGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CATTCCGATTCTGATGAAG	XhoI
687	Forward	CGCGGATCCCATATG-TGCGACAGCAAAGTCCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTGCGCGGCTTTTTGTT	XhoI
690	Forward	CGCGGATCCCATATG-TGTTCTCCGAGCAAAGAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TATTCGCCCCGTGTTTGG	XhoI
691	Forward	CGCGGATCCCATATG-GCCACGGCTTATATCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAGGCAGGAAGAAAG	XhoI

694	Forward	CGCGGATCCCATATG-TTGGTTTCCGCATCCGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCTGCGTCGGTGCGGT	XhoI
695	Forward	CGCGGATCCCATATG-TTGCCTCAAACCTCGTCCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
696	Forward	CGCGGATCCCATATG-TTGGGTTGCCGGCAGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGATTGCCGCAATGATG	XhoI
700a	Forward	AAAAAAGAATTC-GCATCGACAGACGGTGTGTCGTGGAC	Eco RI
	Reverse	AAAAAACTGCAG-TTACGCTACCGGCACGACTTCCAAACC	Pst I
701	Forward	CGCGGATCCCATATG-AAGACTGTTTGGATACTTC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TGCCGACAACAGCCTC	XhoI
702	Forward	AAAAAAGAATTC-ATGCCGTGTTCCAAAGCCAGTTGGATTTC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACCCCATTCACCCGGAGAACCGA	Pst I
703	Forward	CGCGGATCCGCTAGC-CAAACGCTGGCAACCG	BamHI- NheI
	Reverse	CCCGCTCGAG-TTTTGCAGGTTTGATGTTTG	XhoI
704a	Forward	AAAAAAGAATTC-GCTTCTACCGGTACGCTGGCGCG	Eco RI
	Reverse	AAAAAACTGCAG- TTAGTTTTGCCGGATAATATGGCGGGTGCG	Pst I
707	Forward	CGCGGATCCGCTAGC-GAAATTATTAACGATGCAGA	BamHI- NheI
	Reverse	CCCGCTCGAG-GAAACTGTAATTCAAGTTGA	XhoI
708	Forward	CGCGGATCCGCTAGC-CCTTTTAAGCCATCCAAAA	BamHI- NheI
	Reverse	CCCGCTCGAG-TTGACCGGTGAGGACG	XhoI
710	Forward	CGCGGATCCCATATG-GAAACCCACGAAAAAATC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AACGGTTTCGGTCAG	XhoI
714	Forward	CGCGGATCCCATATG-AGCTATCAAGACATCTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCGGTAGGTAAATCGGAT	XhoI
716	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
718	Forward	CGCGGATCCCATATG-GAGCCGATAATGGCAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GGCGCGGGCATGGTCTTGTC	XhoI
720	Forward	CGCGGATCCCATATG-AGCGGATGGCATAACC	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTTGCATAGCTGTTGACCA	XhoI
723	Forward	CGCGGATCCCATATG-CGACCCAAGCCCC	BamHI- NdeI
	Reverse	CCCGCTCGAG-AATGCGAATCCGCCGCC	XhoI

725	Forward	CGCGGATCCCATATG-GTGCGCACGGTTAAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGCTTATCCTTAAGGGTTA	XhoI
726	Forward	CGCGGATCCCATATG-ACCATCTATTTCAAAAAC	BamHI- NdeI
	Reverse	CCCGCTCGAG-GCCGATGTTTAGCGTCC	XhoI
728	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTGAGAAAGGTGCGCG	XhoI
729	Forward	CGCGGATCCCATATG-TGCACCATGATTCCCCA	BamHI- NdeI
	Reverse	GCCCAAGCTT-TTTGTCGGTTTGGGTATC	HindIII
731	Forward	CGCGGATCCGCTAGC-GCCGTGCCGGAGG	BamHI- NheI
	Reverse	CCCGCTCGAG-ACGGGCGCGGCAG	XhoI
732	Forward	CCGGAATTCTACATATG-TCGAAACCTGTTTTTAAGAA	EcoRI- NdeI
	Reverse	CCCGCTCGAG-CTTCTTATCTTTTTTATCTTTC	XhoI
733	Forward	CGCGGATCCCATATG-GCCTGCGGCGGCAA	BamHI- NdeI
	Reverse	CCCGCTCGAG-TCGCTTGCCCTCCTTTAC	XhoI
734	Forward	CGCGGATCCCATATG-GCCGATACTTACGGCTAT	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTTGAGATTTTGAATCAAAGAG	XhoI
735	Forward	CGCGGATCCCATATG-AAGCAGCAGGCGGTCA	BamHI- NdeI
	Reverse	CCCGCTCGAG-ATTTCCGTAGCCGAGGG	XhoI
737	Forward	CGCGGATCCCATATG-CACCACGACGGACACG	BamHI- NdeI
	Reverse	CCCGCTCGAG-GTCGTGCGGCGGGA	XhoI
739	Forward	CGCGGATCCCATATG-GCAAAAAAACCGAACA	BamHI- NdeI
	Reverse	CCCGCTCGAG-GAAGAGTTTGTGCGAGAATT	XhoI
740	Forward	CGCGGATCCCATATG-GCCAATCCGCCCGAAG	BamHI- NdeI
	Reverse	CCCGCTCGAG-AAACGCGCCAAAATAGTG	XhoI
741	Forward	CGCGGATCCCATATG-TGCAGCAGCGGAGGG	BamHI- NdeI
	Reverse	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI
743	Forward	CGCGGATCCCATATG-GACGGTGTGTGCCTGTT	BamHI- NdeI
	Reverse	CCCGCTCGAG-CTTACGGATCAAATTGACG	XhoI
745	Forward	CGCGGATCCCATATG-TTTTGGCAACTGACCG	BamHI- NdeI
	Reverse	CCCGCTCGAG-CAAATCAGATGCCTTTAGG	XhoI
746	Forward	CGCGGATCCCATATG-TCCGAAAACAAACAAAAC	BamHI-

	Reverse	CCCGCTCGAG-TTCATTCGTTACCTGACC	NdeI
747	Forward	CCGGAATTCTAGCTAGC-CTGACCCCTTGGG	XhoI
	Reverse	GCCCAAGCTT-TTTTGATTTTAATTGACTATAGAAC	EcoRI-
749	Forward	CGCGGATCCCATATG-TGCCAGCCGCCG	NheI
	Reverse	CCCGCTCGAG-TTCAAGCCGAGTATGC	HindIII
750	Forward	CGCGGATCCCATATG-TGTCGCCCCGAACCTG	BamHI-
	Reverse	CCCGCTCGAG-CTTTTCCCCGCCGCAA	NdeI
758	Forward	CGCGGATCCCATATG-AACAATCTGACCGTGTT	XhoI
	Reverse	CCCGCTCGAG-TGGCTCAATCCTTTCTGC	BamHI-
759	Forward	CGCGGATCCGCTAGC-CGCTTCACACACACCAC	NdeI
	Reverse	CCCGCTCGAG-CCAGTTGTAGCCTATTTTG	XhoI
763	Forward	CGCGGATCCCATATG-CTGCCTGAAGCATGGCG	BamHI-
	Reverse	CCCGCTCGAG-TTCCGCAAATACCGTTTCC	NdeI
764	Forward	CGCGGATCCCATATG-TTTTCTCCGCCCTGA	XhoI
	Reverse	CCCGCTCGAG-TCGCTCCCTAAAGCTTTC	BamHI-
765	Forward	CGCGGATCCCATATG-TTAAGATGCCGTCCG	NdeI
	Reverse	CCCGCTCGAG-ACGCCGACGTTTTTTATTAA	XhoI
767	Forward	CGCGGATCCCATATG-CTGACGGAAGGGGAAG	BamHI-
	Reverse	CCCGCTCGAG-TTTCTGTACAGCAGGGG	NdeI
768	Forward	CGCGGATCCCATATG-GCCCCGAAAAACCCG	XhoI
	Reverse	CCCGCTCGAG-TTTCATCCCTTTTTTGAGC	BamHI-
770	Forward	CGCGGATCCCATATG-TGCGGCAGCGGCGAA	NdeI
	Reverse	CCCGCTCGAG-GCGTTTGTGCGAGATTTTC	XhoI
771	Forward	CGCGGATCCCATATG-TCCGTATATCGCACCTTC	BamHI-
	Reverse	CCCGCTCGAG-CGGTTCTTTAGGTTTGAG	NdeI
772	Forward	CGCGGATCCCATATG-TTGCGGCGTTGGTGG	XhoI
	Reverse	CCCGCTCGAG-CAATGCCGACATCAAACG	BamHI-
774	Forward	CGCGGATCCCATATG-TCCGTTTACCCGTTCC	NdeI
	Reverse	CCCGCTCGAG-TCGTTTGCGCACGGCT	XhoI
790	Forward	CGCGGATCCCATATG-GCAAGAAGGTCAAAAAC	BamHI-
			NdeI

	Reverse CCCGCTCGAG-GGCGTTGTTCCGATTTCG	XhoI
900	Forward CGCGGATCCCATATG-CCGTCTGAAATGCCG	BamHI- NdeI
	Reverse CCCGCTCGAG-ATATGAAAAGTCTGTTGTC	XhoI
901	Forward CGCGGATCCCATATG-CCCGATTTTTTCGATG	BamHI- NdeI
	Reverse CCCGCTCGAG-AAAATGGAACAATACCAGG	XhoI
902	Forward. CCGGAATTCTACATATG-TTGCACTTTCAAAGGATAATC	EcoRI- NdeI
	Reverse CCCGCTCGAG-AAAAATGTACAATGGCGTAC	XhoI
903	Forward CCGGAATTCTAGCTAGC-CAGCGTCAGCAGCACAT	EcoRI- NheI
	Reverse CCCGCTCGAG-GAAACTGTAATTCAAGTTGAA	XhoI
904	Forward AAAAAAGGTACC-ATGATGCAGCACAATCGTTTC	Kpn I
	Reverse AAAGTGCAG-TTAATATCGATAGGTTATATG	Pst I
904a	Forward AAAAAAGAATTC-CGGCTCGGCATTGTGCAGATGTTGCA	Eco RI
	Reverse AAAGTGCAG-TTAATATCGATAGGTTATATG	Pst I
905	Forward CGCGGATCCCATATG-AACAAAATATACCGCATC	BamHI- NdeI
	Reverse CCCGCTCGAG-CCACTGATAACCGACAGAT	XhoI
907	Forward CGCGGATCCCATATG-GGCGCGCAACGTGAG	BamHI- NdeI
	Reverse CCCGCTCGAG-ACGCCACTGCCAGCG	XhoI
908	Forward AAAGAATTC-GCAGAGTTAGTAGGCGTTAATAAAAATAC	Eco RI
	Reverse AAAGTGCAG-TTAATATGGTTTTGTCGTTTCG	Pst I
909	Forward CGCGGATCCCATATG-TGCGCGTGGGAAACTTAT	BamHI- NdeI
	Reverse CCCGCTCGAG-TCGGTTTTGAACTTTGGTTTT	XhoI
910	Forward AAAGAATTC-GCATTGCGCGGCGACTCTGCCGAGCG	Eco RI
	Reverse AAAGTGCAG-TCAGCGATCGAGCTGCTCTTT	Pst I
911	Forward AAAGAATTC-GCTTTCCGCGTGGCCGGCGGTGC	Eco RI
	Reverse AAAAAACTGCAG-GTCGACTTATTCGGCGGCTTTTCCGC	Pst I
912	Forward AAAAAAGAATTC- CAATCCGTCAAAACGCCACTCAAGTATTGAG	Eco RI
	Reverse AAAAAACTGCAG-TTACAGTCCGTCCACGCCTTTCGC	Pst I
913	Forward CGCGGATCCCATATG-GAAACCCGCCCCGC	BamHI- NdeI
	Reverse CCCGCTCGAG-AGGTTGTGTTCCAGGTTG	XhoI
915	Forward CGCGGATCCCATATG-TGCCGGCAGGCGGAA	BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGAAAATATAGGTATCAGG	XhoI
914	Forward AAAGAATTC-GACAGAATCGGCGATTTGGAAGCACG	Eco RI
	Reverse AAAGTGCAG-CTATATGCGCGGCAGGACGCTCAACGG	Pst I
916	Forward CGCGGATCCCATATG-GCAATGATGGCGGCTG	BamHI- NdeI
	Reverse CCCGCTCGAG-TTTGGCGGCATCTTTCAT	XhoI

917	Forward	AAAAAAGAATTC-CCTGCCGAAAAACCGGCACCGGC	Eco RI
	Reverse	AAAAAACTGCAG-TTATTTCCCGCCTTCACATCCTG	Pst I
919	Forward	CGCGGATCCCATATG-TGCCAAAGCAAGAGCATC	BamHI-
	Reverse	CCCGCTCGAG-CGGGCGGTATTTCGGG	NdeI
920	Forward	CGCGGATCCCATATG-CACCGCGTCTGGGTC	XhoI
	Reverse	CCCGCTCGAG-ATGGTGCGAATGACCGA	BamHI-
921	Forward	AAAAAAGAATTC-TTGACGGAAATCCCGTGAATCC	NdeI
	Reverse	AAAAAACTGCAG-TCATTTCAAGGGCTGCATCTTCAT	XhoI
922	Forward	CGCGGATCCGCTAGC-TGTACGGCGATGGAGGC	Eco RI
2	Reverse	CCCGCTCGAG-CAATCCCGGGCCGCC	Pst I
923	Forward	CGCGGATCCCATATG-TGTTACGCAATATTGTCCC	BamHI-
	Reverse	CCCGCTCGAG-GGACAAGGCGACGAAG	NheI
925	Forward	CGCGGATCCCATATG-AAACAAATGCTTTTAGCCG	XhoI
	Reverse	CCCGCTCGAG-GCCGTTGCATTTGATTTT	BamHI-
926	Forward	CGCGGATCCCATATG-TGCGCGCAATTACCTC	NdeI
	Reverse	CCCGCTCGAG-TCTCGTGCGCGCCG	XhoI
927	Forward	CGCGGATCCCATATG-TGCAGCCCCGCAGC	BamHI-
	Reverse	CCCGCTCGAG-GTTTTTTGCTGACGTAGT	NdeI
929a	Forward	AAAAAAGAATTC-CGCGGTTTGCTCAAAACAGGGCTGGG	XhoI
	Reverse	AAAAAATCTAGA-TTAAGAAAGACGGAACTACTGCC	Eco RI
931	Forward	AAAAAAGAATTC-GCAACCCATGTTTTGATGGAAAC	Xba I
	Reverse	AAAAAACTGCAG-TTACTGCCCGACAACAACGCGACG	Eco RI
935	Forward	AAAAAAGAATTC-GCGGATGCGCCCGCGATTTTGGATGACAAGGC	Pst I
	Reverse	AAAAAACTGCAG-TCAAAACCGCCAATCCGCCGACAC	Eco RI
936	Forward	CGCGGATCCCATATG-GCCGCCGTCGGCGC	Pst I
	Reverse	CCCGCTCGAG-GCGTTGGACGTAGTTTTG	BamHI-
937	Forward	AAAAAAGAATTC-CCGGTTTACATTCAAACCGGCGCAAC	NdeI
	Reverse	AAAAAACTGCAG-TTAAAATGTATGCTGTACGCCAAA	XhoI
939a	Forward	AAAAAAGAATTC-GGTTTCGGCAGCTGTGATGAAACC	Eco RI
	Reverse	AAAAAACTGCAG-TTAACGCAAACCTTGGATAAAGTTGGC	Pst I
950	Forward	CGCGGATCCCATATG-GCCAACAAACCGGCAAG	BamHI-
	Reverse	CCCGCTCGAG-TTAGAACCGCATTTGCC	NdeI
953	Forward	CGCGGATCCCATATG-GCCACCTACAAAGTGGAC	XhoI
	Reverse	CCCGCTCGAG-TTGTTTGGCTGCCTCGAT	BamHI-
957	Forward	CGCGGATCCCATATG-TTTTGGCTGGGAACGGG	NdeI
			XhoI
			BamHI-

	Reverse CCCGCTCGAG-GTGAGAAAGGTCGCGC	NdeI
958	Forward CGCGGATCCCATATG-GCCGATGCCGTTGCG	XhoI
		BamHI-
	Reverse GCCCAAGCTT-GGGTCGTTTGTGTCGTC	NdeI
959	Forward CGCGGATCCCATATG-CACCACGACGGACACG	HindIII
		BamHI-
	Reverse CCCGCTCGAG-GTCGTCGCGGCGGGA	NdeI
961	Forward CGCGGATCCCATATG-GCCACAAGCGACGACG	XhoI
		BamHI-
	Reverse CCCGCTCGAG-CCACTCGTAATTGACGC	NdeI
972	Forward AAAAAAGAATTC- TTGACTAACAGGGGGGGAGCGAAATTAAAAAC	XhoI
		Eco RI
	Reverse AAAAAATCTAGA-TTAAAAATAATCATAATCTACATTTTG	Xba I
973	Forward AAAAAAGAATTC-ATGGACGGCGCACAAACCGAAAAC	Eco RI
	Reverse AAAAAACTGCAG-TTACTTCACGCGGGTCGCCATCAGCGT	Pst I
982	Forward CGCGGATCCCATATG-GCAGCAAAGACGTAC	BamHI-
		NdeI
	Reverse CCCGCTCGAG-CATCATGCCGCCCATCC	XhoI
983	Forward CGCGGATCCCATATG-TTAGCTGTTGCAACAACAC	BamHI-
		NdeI
	Reverse CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
987	Forward CGCGGATCCCATATG-CCCCACTGGAAGAAC	BamHI-
		NdeI
	Reverse CCCGCTCGAG-TAATAAACCTTCTATGGGC	XhoI
988	Forward CGCGGATCCCATATG-TCTTTAAATTTACGGGAAAAAG	BamHI-
		NdeI
	Reverse GCCCAAGCTT-TGATTTGCCTTTCCGTTTT	HindIII
989	Forward CCGGAATTCTACATATG-GTCCACGCATCCGGCTA	EcoRI-
		NdeI
	Reverse CCCGCTCGAG-TTTGAATTTGTAGGTGTATTGC	XhoI
990	Forward CGCGGATCCGCTAGC-TTCAGAGCTCAGCTT	BamHI-
	2	NheI
	Reverse CCCGCTCGAG-AAACAGCCATTGAGCGA	XhoI
992	Forward CGCGGATCCCATATG-GACGCGCCCGCCCG	BamHI-
		NdeI
	Reverse CCCGCTCGAG-CCAAATGCCCAACCATTC	XhoI
993	Forward CGCGGATCCCATATG-GCAATGCTGATTGAAATCA	BamHI-
		NdeI
	Reverse CCCGCTCGAG-GAACACATCGCGCCCG	XhoI
996	Forward CGCGGATCCCATATG-TGCGGCAGAAAATCCGC	BamHI-
		NdeI
	Reverse CCCGCTCGAG-TCTAAACCCCTGTTTCTC	XhoI
997	Forward CCGGAATTCTAGCTAGC-CGGCACGCCGACGTT	EcoRI-
		NheI
	Reverse CCCGCTCGAG-GACGGCATCGCTCAGG	XhoI

Underlined sequences indicate restriction recognition sites.

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 1>:

```
g001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG GTGTCGGCGA ACGAGGTGTC
51  CGGCAGGGCT TCGCCCCGA TGGTGCTGGT CATCTGCCAG ACGCTGCCGA
101 AACGCGATAC TTAAACGGC TCGGGTACGC ATACTTTACC GGTTTGGGCG
151 ATTTTGCCGA GGTCGTGCG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCGGTTT GTAACTCGGC GCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCCGT GCTGATGGTG
301 CCGTCTGAAG CGATGTTGAG GAAGAGTTCG GCGGAGAAAC ACAGCGTCCA
351 CCGGATTGTC CCGGCTTCAT CGGGAGGTG GGACAATACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 2; ORF 001.ng>:

```
g001.pep
1  MLPOGKAARR VSANEVSGRA CARMVLVICQ TLPKRDTLNG SGTHTLPVWA
51  ILPRSLRSKS TIITFSARFF GSVNSAARR SSCPSPKIGA VFFIGSVLMV
101 PSEAMLRKSS GEKHSVHAD C PASSGRWDNT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 3>:

```
m001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTG
51  CGGcAssCTT ss.GCTTGA yGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
151 ATTTTGCCGA GATCGTTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCCGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GCGGAGAAAC ACAGCGTCCA
351 CCGGATTGTC CCCTCCGCAT CGGGAGGTG GGACAAGACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 4; ORF 001>:

```
m001.pep
1  MLPOGKAARR MSANEVCGXL XAWXVLVICQ TLPKRDTLNG SGTHTVPVWA
51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VFFIGSVLMV
101 PSEPILRKSS GEKHSVHAD C PSASGRWDKT A*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 5>:

```
a001.seq
1  ATGCTGCCGC AGGGGAAGGC GCGCGGAGG ATGTCGGCGA ACGAGGTGTG
51  CGGCAAGGCT TGGGCTTGA TGGTGCTGGT CATCTGCCAA ACGCTGCCGA
101 AACGCGATAC TTAAACGGT TCGGGTACGC ATACTGTGCC GGTTTGGGCG
151 ATTTTGCCGA GGTCGTACG CAGCAAATCG ACAATCATCA CGTTTTCGGC
201 GCGGTTTTTC GGGTCTGCTT GCAACTCGGC GCGCGGCGT TCGTCTTGTC
251 CGTCGCCCAA AATCGGCGCG GTGCCTTTCA TCGGTTCCGT GCTGATGGTG
301 CCGTCCGAAC CGATTTTGAG GAAGAGTTCG GCGGAGAAAC ACAGCGTCCA
351 CCGGATTGTC CCTTGTGCAT CGGGAGGTG GGACAAAACG GCATAG
```

This corresponds to the amino acid sequence <SEQ ID 6; ORF 001.a>:

```
a001.pep
1  MLPOGKAARR MSANEVCGKA WAWMVLVICQ TLPKRDTLNG SGTHTVPVWA
51  ILPRSLRSKS TIITFSARFF GSACNSAARR SSCPSPKIGA VFFIGSVLMV
101 PSEPILRKSS GEKHSVHAD C PCASGRWDKT A*
```

m001/a001 96.2% identity over a 131 aa overlap

	10	20	30	40	50	60
m001.pep	MLPOGKAARRMSANEVCGXLXAWXVLVICQ	TLPKRDTLNGSGTHTVPVWAILPRSLRSKS				
a001.pep	MLPOGKAARRMSANEVCGKAWAWMVLVICQ	TLPKRDTLNGSGTHTVPVWAILPRSLRSKS				

	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
a001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
	70	80	90	100	110	120
	130					
m001.pep	PSASGRWDKTAX					
a001.pep	PCASGRWDKTAX					
	130					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 001 shows 89.3% identity over a 131 aa overlap with a predicted ORF (ORF 001.ng) from *N. gonorrhoeae*:

m001/g001

	10	20	30	40	50	60
m001.pep	MLPQGKAARRMSANEVCGXLXAWXVLVICQTLPKRDTLNGSGTHTVPVWAILPRSLRSKS					
g001	MLPQGKAARRVSANEVSGRACARMVLVICQTLPKRDTLNGSGTHTLPVWAILPRSLRSKS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m001.pep	TIIITFSARFFGSACNSAARRSSCPSPKIGAVPFIGSVLMVPSEPILRKSSGEKHSVHADC					
g001	TIIITFSAREFGSVCNSAARRSSCPSPKIGAVPFIGSVLMVPSEAMLRKSSGEKHSVHADC					
	70	80	90	100	110	120
	130					
m001.pep	PSASGRWDKTAX					
g001	PASSGRWDNTAX					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 7>:

g003.seq

```

1  ATGGTCGTAT TCGTGGCTGA AGGCGTATTC GGTGCGCTG TTTTGGGTCA
51  CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCCGC GTCAC TCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGCTTTGGT
151 TTTGCCCGGC AGCGGTTTCGT CGGCTTTGCG GATGTCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAGG TAGTCCTGAT GGTATTCCTC GGCGTCGTAG
251 AAGTTTTTCA GCGGTTTCGT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTTC GCGGGGTCGG
351 TGTAGTACAC GCCGCTGCGG TATTGCGTGC CGGTGTCGTT ACCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAA TAATATTGCA GGATGTCGTC
451 CAGGCTGagt TTGTCGGCAT CTaggtcac tTTGACGGTC TCGGCATGAC
501 CCGTATGGCG GTaggacact tctTCgtanc TcGGGtTTTC CGTgttGCCG
551 TTGGCgttac cGGATACCGC gtcaACCACG CCGTcgatgc gttggaAATa
601 ggCTTCCAAG ccccaaaagc agccgccggc gaagtaaatg gtgcccgtgt
651 tcatgattGC TGa

```

This corresponds to the amino acid sequence <SEQ ID 8; ORF 003.ng>:

g003.pep

```

1  MVVFVAEGVF GRAVLGHLVL LFGQGAFEFG VTRFFIRCRV EAFALRCGFG
51  FARQRFVGFA DVDVAVAVGV FNQVVLMLVFL GVVEVFQRFV FNNEGQLVFL
101 LLAFEGGGDD GFFGGVGVVH AAVALRAGVV TLFVEAGRIN DAEIILQDVV

```

151 QAEFVGIVGH FDGLGMTRMA VGHFFVRVFR VAVGVTGYRV NHAVDAL EIG
201 FQAPKAAAGE VNGARVHDC

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 9>:

m003.seq
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTGCGCGTG TTTTGGGTAA
51 CTTGsTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACCTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGGG CGGTCTTGGT
151 TTTGCCCGGC AGCGGTTCGT CAGCKTTGCG GATGTCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251 AAGTTTTtCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301 CTGCTCGCGT TTGAGGGCGK CGGCGATGAC GGCTTTTTTCG kCGGGGTCGG
351 TGATGTACAC GCCGCTGCGG TATTGCGTAC CGGTGTCGTT GCCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTGCGCG
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAATA
601 GGCTTCCAAG CCCAGAAGC AGCg.CCGGC GAGGTAAATG GTGCGCGTGT
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 10; ORF 003>:

m003.pep Length: 221
1 MVVFVAEGIF GRAVLGNLXL LFGQGAFFFG VTRFFIRCRV EAFALRGGLG
51 FARQRFVSXA DVDVAVAVGV FNQVLMVFL GIVEVFQRLV FNNEGQLVFL
101 LLAFEGXGDD GFFXGVGVVH AAVALRTGVV ALFVEAGRIN DAEEILQDGV
151 *AEFVGIVGH FDGFGVARMA VGHVFIARIF RAVGVAGYR VNHAVDAL E I
201 GFQAPEAAXG EVNGARVHDF *

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 11>:

a003.seq
1 ATGGTCGTAT TCGTGGCTGA AGGCATATTC GGTGCGCGTG TTTTGGGTAA
51 CTTGGTATTG CTCTTCGGTC AGGGTGCGTT TGAGTTCGGC GTCACCTCGGT
101 TTTTATACG TTGCCGCGTC GAAGCCTTTG CCTTGCGGTG CGGTCTTGGT
151 TTTGCCCGGC AGCGGTTCGT CGGCTTTGCG GATATCGATG TGGCAGTAGC
201 CGTTGGGGTT TTTAATCAAG TAGTCCTGAT GGTATTCCTC GGCATCGTAG
251 AAGTTTTTCA GCGGCTCGTT TTCAACAACG AGGGGCAGTT GGTATTTTTG
301 CTGCTCGCGT TTGAGGGCGG CGGCGATGAC GGCTTTTTTCG GCGGGGTCGG
351 TGATGTACAC GCCGCTGCGG TATTGCGTAC CGGTGTCGTT GCCCTGTTTG
401 TTGAGGCTGG TCGGATCAAC GACGCGGAAG AAATATTGCA GGATGTCGTC
451 TAGGCTGAGT TTGTCGGCAT CGTAGGTCAC TTTGACGGTT TCGGCGTGGC
501 CCGTATGGCG GTAGGACACG TCTTCATAGC TCGGATTTTT CGTGTGCGCG
551 TTGGCGTAGC CGGATACCGC GTCAACCACG CCGTCGATGC GTTGGAATA
601 GGCTTCCAAG CCCAGAAGC AGCCGCCGGC GAGGTAGATG GTGCGCGTGT
651 TCATGATTTT TGA

This corresponds to the amino acid sequence <SEQ ID 12; ORF 003.a>:

a003.pep
1 MVVFVAEGIF GRAVLGNLVL LFGQGAFFFG VTRFFIRCRV EAFALRCGLG
51 FARQRFVGFA DIDVAVAVGV FNQVLMVFL GIVEVFQRLV FNNEGQLVFL
101 LLAFEGGGDD GFFGGGVGVH AAVALRTGVV ALFVEAGRIN DAEEILQDGV
151 *AEFVGIVGH FDGFGVARMA VGHVFIARIF RAVGVAGYR VNHAVDAL E I
201 GFQAPEAAG EVNGARVHDF *

m003/a003 95.9% identity over a 220 aa overlap

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVLGNLXL	LFGQGAFFFGVTRFFIRCRV	EAFALRGGLGFARQRFVSXA			
a003	MVVFVAEGIFGRAVLGNLVL	LFGQGAFFFGVTRFFIRCRV	EAFALRCGLGFARQRFVGFA			
	10	20	30	40	50	60

168

	70	80	90	100	110	120
m003.pep	DVDVAVAVGVFNQVVLVFLGIVEVFQRLVFNNEGQLVFLLLAFEGXGDDGFFXGVGVVH					
a003	DIDVAVAVGVFNQVVLVFLGIVEVFQRLVFNNEGQLVFLLLAFEGGGDDGFFGGVGVVH					
	70	80	90	100	110	120
	130	140	150	160	170	180
m003.pep	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAGHVFIARIF					
a003	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAGHVFIARIF					
	130	140	150	160	170	180
	190	200	210	220		
m003.pep	RVAVGVAGYRVNHAVDALFQAPAAAXGEVNGARVHDFX					
a003	RVAVGVAGYRVNHAVDALFQAPAAAGEVDGARVHDFX					
	190	200	210	220		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 003 shows 88.6% identity over a 219 aa overlap with a predicted ORF (ORF 003.ng) from *N. gonorrhoeae*:

m003/g003

	10	20	30	40	50	60
m003.pep	MVVFVAEGIFGRAVLGNLXLLFGQGAFFFGVTRFFIRCRVEAFALRGGLGFARQRFVVSXA					
g003	MVVFVAEGVFGRAVLGHLVLLFGQGAFFFGVTRFFIRCRVEAFALRCGFGFARQRFVGF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m003.pep	DVDVAVAVGVFNQVVLVFLGIVEVFQRLVFNNEGQLVFLLLAFEGXGDDGFFXGVGVVH					
g003	DVDVAVAVGVFNQVVLVFLGVVEVFQRFVFNNEGQLVFLLLAFEGGGDDGFFGGVGVVH					
	70	80	90	100	110	120
	130	140	150	160	170	180
m003.pep	AAAVLRTGVVALFVEAGRINDAEEILQDVVXAEFVGIVGHFDGFGVARMAGHVFIARIF					
g003	AAAVLRAGVTLFVEAGRINDAEEILQDVVQAEFVGIVGHFDGLGMTRMAVGHFFV-RVF					
	130	140	150	160	170	180
	190	200	210	220		
m003.pep	RVAVGVAGYRVNHAVDALFQAPAAAXGEVNGARVHDFX					
g003	RVAVGVGTGYRVNHAVDALFQAPKAAAGEVNGARVHDC					
	190	200	210			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 13>:

g004.seq

```

1  ATGgtagAAC GGCATATCCA GCATTGCGG AACGGTCATC TTCATTGAT
51  GCGCCCATGC CAACAagtga gccAAatgtT CGGCGGCAGG GCCTacgatT
101 TCCGCGCCGA TAAagcggcc gGTGgctTTT tCGGCataca ggcgcaTatg
151 gCCTTTGTTT ACCAgcatca cgcggtcgcg accttgatTT TTGAACGATA
201 CTTCGCCgaT GACAAATTCG TCGGCTTGGT ATTGCGCGCG AACCTGCGCG
251 TATTTCAAAC CGACAAAGCC GATTTCGgga ctggttaaACA CCACGCCAAT
301 GGTgctgcgg cGCAAACCGC TGCCGATATt cgGtagcgg ccccgcggtta
351 ttgcccggca atcttacctt ggtcggcggc ttcatGCAGC AGGGGcagtt
401 ggtttggacgc gtcgcccgcg ataAAGATAT GCGGAATgct ggtCTGCATg
451 gtCAGCGGAT CGGCAACGGG tacgcccgcg gcgtctttgT CGATATTGAT
501 GTTTTCCAAA CCGATATtgT CAACGTTCGG ACGGCgACCT ACGGCTGCCA

```

```

551 ACATATATTC GGCAACAAAT ACGCCTTTTT CGCCATCCTG CTCCCAATGG
601 ACTtcttACAT TGCCGTCTGC GTCGAGTTTG ACCTCGGTTT TAGCATCCAG
651 ATGCAGTTTC AATtcttTCTC CGAACACGGC TTTCGCCTCG TCTGAAACAA
701 CGGGGTCCGA AATGCCGCCG ATGATTCCGC CCAAACCGAA AATTTCAACT
751 TTCACACCCA AACGGTGCAA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 14; ORF 004.ng>:

```

g004.pep
1 MVERHIQHLR NGHLLMRPC QQVSQMFGR AYDFRADKAA GGFFGIQAHM
51 AFVYQHHAHA TLIFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAN
101 GAAATTAADI RVAAPRYCPA ILPWSAASCS RGSWLDASPA IKICGMLVCM
151 VSGSATGTFR ASLSILMFSK PILSTFGRRP TAANIYSATN TPFSPSCSQW
201 TSTLPSASSL TSVLASRCSF NSSPNTAFAS SETTGSEMPF MIPPKPKIST
251 FTPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 15>:

```

m004.seq
1 ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTGAT
51 GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCAGG GCCTACGATT
101 TCCGCGCCGA TAAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTC ACCAGCATCA CGCGGCTGCG GCCTTGTTT TTGAACGATA
201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTAGAC CGACAAAGCC GATTTCGCGA CTGGTAAACA CCACGCCGAT
301 GGTGCTGCGC CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGGCAATC TTGCCTTGGT CGGCAGCTTC ATGCAGCAGA GGCAGTTGGT
401 TGGACGCATC GCCTGCGATG AAGATATGCG GAATACTGGT CTGCATGGTC
451 AGCGGGTCGG CAACAGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATATT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCATC GAGTTTGACC TCGGTTTtag CATCCAGATG
651 CAGTTTCAAT TCTTCGCCGA ACACGGCGTT CGCCTCGTCT GAAACGACGG
701 GGTCCGAAAT GCCGCCGATG ATTCCGCCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCAAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 16; ORF 004>:

```

m004.pep
1 MVERHIQHLR NGHLLMCPs QQVROMFGR AYDFRADKAA GGFFGIQAHM
51 AEFVQHHAHA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGKHHAD
101 GAAPQTAADI RVAAALSPAI LPWSAASCSR GSWLDASPA KICGILVCMV
151 SGSATGTFR SFSILIFSKP ILSTFGRRP AASIYSATNT PFSPPSCSQW
201 STLPSASSLT SVLASRCSFN SSPNTAFAS ETTGSEMPF IPPPKISTF
251 TPKRCNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 17>:

```

a004.seq
1 ATGGTAGAAC GGCATATCCA GCATTTGCGG AACGGTCATC TTCATTGAT
51 GTGCCCAAGC CAACAGGTGC GCCAAATGTT CGGCGGCCG ACCTACGATT
101 TCTGCGCCGA TGAAGCGGCC GGTGGCTTTT TCGGCATACA GGCGCATATG
151 GCCTTTGTTT ACCAGCATCA CGCGGCTGCG GCCTTGTTT TTGAACGATA
201 CTTCGCCGAT GACAAATTCG TCGGCTTGGT ATTGCGCGGC AACCTGCGCG
251 TATTTCAAAC CGACAAAGCC GATTTCGCGA CTGGTGAACA CTACGCCGAT
301 GGTGCTGCGG CGCAAACCGC CGCCGATATT CGGGTAGCGG CCGCGTTATC
351 GCCGGCAATC TTGCCTTGGT CGGCGGCTTC ATGCAGCAGG GGCAGTTGGT
401 TGGACGCGTC GCCCGAATA AAGATATGCG GAATACTGGT CTGCATAGTC
451 AGCGGATCGG CAACGGGTAC GCCGCGCGCA TCTTTTTCGA TATTGATGT
501 TTCCAAACCG ATATTGTCAA CGTTCGGACG GCGGCCTACG GCTGCCAGCA
551 TATATTCGGC AACAAATACG CCTTTTTCGC CATCCTGCTC CCAATGGACT
601 TCTACATTGC CGTCTGCGTC GAGTTTGCC TCGGTTTtag CATCCAAATG
651 CAGTTTCAAT TCTTCACCGA ACACGGCTTT CGCCTCGTCT GAAACGACGG
701 GGTCCGAAAT GCCGCCGATG ATGCCACCCA AACCGAAAAT TTCAACTTTC
751 ACGCCCAAAC GGTGCAATGC CTGA

```

This corresponds to the amino acid sequence <SEQ ID 18; ORF 004.a>:

```
a004.pep
1  MVERHIQHLR NGHLHLMCPS QQVRQMFGGR TYDFCADEAA GGFFGIQAHM
51 AFVYQHAAA ALVFERYFAD DKFVGLVLRG NLRVFQTDKA DLRTGEHYAD
101 GAAQAATAADI RVAAALSPAI LPWSAASCSR GSWLDASPAI KICGILVCIV
151 SGSATGTPRA SFSILMFSKP ILSTFGRRT AASIYSATNT PFSPPSCSQWT
201 STLPSASSLA SVLASKCSFN SSPNTAFASS ETTGSEMPPM MPPKPKISTF
251 TPKRCNA*
```

m004/a004 94.9% identity over a 257 aa overlap

```

      10      20      30      40      50      60
m004.pep MVERHIQHLRNGHLHLMCPSQQVRQMFGGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a004      MVERHIQHLRNGHLHLMCPSQQVRQMFGGRTYDFCADEAAGGFFGIQAHMAFVYQHAAA
          10      20      30      40      50      60

      70      80      90      100     110     120
m004.pep ALVFERYFADDKFVGLVLRGNLRFQTDKADLRTGKHHADGAAPQTAADIRVAAALSPAI
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a004      ALVFERYFADDKFVGLVLRGNLRFQTDKADLRTGEHYADGAAQAATAADIRVAAALSPAI
          70      80      90      100     110     120

      130     140     150     160     170     180
m004.pep LPWSAASCSRGSWLDASPAIKICGILVCMVSGSATGTPRASFSILIFSILSKPILSTFGRRT
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a004      LPWSAASCSRGSWLDASPAIKICGILVCIVSGSATGTPRASFSILMFSKPIILSTFGRRT
          130     140     150     160     170     180

      190     200     210     220     230     240
m004.pep AASIYSATNTPFSPPSCSQWTSTLPSASSLTVLASRCSFNSSPNTAFASSETTGSEMPPM
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a004      AASIYSATNTPFSPPSCSQWTSTLPSASSLTVLASRCSFNSSPNTAFASSETTGSEMPPM
          190     200     210     220     230     240

      250
m004.pep IPPKPKISTFTPKRCNAX
          :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a004      MPPKPKISTFTPKRCNAX
          250
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 004 shows 93.4% identity over a 258 aa overlap with a predicted ORF (ORF 004.ng) from *N. gonorrhoeae*:

m004/g004

```

      10      20      30      40      50      60
m004.pep MVERHIQHLRNGHLHLMCPSQQVRQMFGGRAYDFRADKAAGGFFGIQAHMAFVHQHHAAA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g004      MVERHIQHLRNGHLHLMRQCQVSMFGGRAYDFRADKAAGGFFGIQAHMAFVYQHAAA
          10      20      30      40      50      60

      70      80      90      100     110     119
m004.pep ALVFERYFADDKFVGLVLRGNLRFQTDKADLRTGKHHADGAAPQTAADIRVAAA-LSPA
          :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g004      TLIFERYFADDKFVGLVLRGNLRFQTDKADLRTGKHHANGAAAQTAADIRVAAPRYCPA
          70      80      90      100     110     120

      120     130     140     150     160     170     179
m004.pep ILPWSAASCSRGSWLDASPAIKICGILVCMVSGSATGTPRASFSILIFSILSKPILSTFGRRT
```


	: :
g004	ILPWSAASCSCRGSWLDASPAIKICGMLVCMVSGSATGTTPRASLSILMFSPKIPILSTFGRRP
	130 140 150 160 170 180
m004.pep	180 190 200 210 220 230 239
	TAASIYSATNTPFSPSCSQWSTLPSASSLTSVLASRCSFNSSPNPTAFASSETTGSEMPP
g004	:
	TAANIYSATNTPFSPSCSQWSTLPSASSLTSVLASRCSFNSSPNPTAFASSETTGSEMPP
	190 200 210 220 230 240
m004.pep	240 250
	MIPPKPKISTFTPKRCNAX
g004	
	MIPPKPKISTFTPKRCNA
	250

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 19>:

```

g005.seq
1  ATGGGGATGG ACAATATTGA TATGTTTCATG CCTGAACAAG AGGAAATCCA
51  ATCAATGTGG AAAGAAATTT TACTGAATTA CGGTATTTTC CTGCTCGAAC
101 TGCTTACCGT GTTTCGCGCA TACTGCGTGA TTGTGTTGGC TATCGTACAG
151 AGTAAGAAAC AGTCGGAAAG CGGCAGTGTC GTACTGACAG ATTCTTCGGA
201 AAATTATAAA AAACAGCGGC AATCGTTTGA AACATTCTTT TTAAGCGAGG
251 AAGACACAAA ACATCAGGAA AAAAAAGAAA AGAAAAAGGA AAAGCGCGAA
301 GCGAAAGCAG AGAAAAAGCG TTGAAGGAG GCGCGGGAGA AATCTGCCGA
351 AACGCAAAAA TCCCGCCTTT TTGTGTTGGA TTTTGACGGC GATTGTTATG
401 CACACGCCGT AGAATCCTTG CGTCATGAGA TTACGGCGGT GCTTTTGATT
451 GCCAAGCCTG AAGATGAGGT TGTGCTCAGA TTGAAAAGTC CGGGCGCGCT
501 GGTTCACGGT TACG GTTTGG CCGCTTCGCA AGCTAGGCGT TTGCGCGAAT
551 GCAATATTCC GCTGAccgtc gccgTCGATA GGGTCGCGGC AAGCGcgcg
601 tatatgatgg cgtgtgtgGC GGATAAAATT GTTCCGCTc cgtttgcggt
651 catcggttcg gtggggtgtg tggcggaagt gcCGAATATC CaccgCctGT
701 TGAAAAAACA TGATATTGAT GTGGATGTGA TGACGCGCGG CGAATTTAAG
751 CGCACGGTTA CTTTTATGGG TGAAAAATACG GAAAAAGGCA AACAGAAATT
801 CCGGACGGAA CTGGAGGAAA CGCATCAGTT GTTCAAGCAG TTTGTCAGTG
851 AAAACCGCCC CGGGTTGGAT ATTGAAAAAA TAGCGACGGG CGAGCATTTG
901 TTGCGCGCGC AGGCGTTGGC GTTGAAC TTGACGGA ATTGACGAGG CTGCGACCA
951 TGATGATTTG TTGTTGAAAG CTTTGA AAAACAGGgt atCGAAGTGA
1001 AATATCAGGA GAAGCGAAGC CTGATCCAGC GCATTGGTTT GCAGGCGGAA
1051 GCTTCCGTTG AAAAGTTGTT TGCCAAACTT GTCAACCGGC GAGCGGATGT
1101 GATGTAG

```

This corresponds to the amino acid sequence <SEQ ID 20; ORF 005.ng>:

```

g005.pap
1  MGDMDNIDMF  PEQEEIQSMW  KEILLNYGIF  LLELLTVFGA  IALIVLAIQV
51  SKKQSESGSV  VLTDFSENY  KQRQSFETFF  LSEEETHKQE  KKEKKKKEAE
101 AKAEKKRLKE  GGEKSAETQK  SRLFVLDFDG  DLYAHAVESL  RHEITAVILL
151 AKPEDEVLLR  LESPGGVVHG  VGLAASQLRR  LERNIPLTV  AVDKVAASGG
201 YMMACVADKI  VSAPFAVIGS  VGVVAEVPNI  HRLKKKHDI  DVDMTAGEFK
251 RTVTFMGENT  EKGKQKFRQE  LEETHQLFKQ  FVSENRPGLD  IEKIATGEHW
301 FGRQALALNL  IDEISTDDLL  LLKAFENKQV  IEVKYQEKRS  LIQRIGLQAE
351 ASVEKFLAKL  VNRRADVM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEO ID 21>:

```
m005.seq
1 ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51 GTGGAAGAA ATTTTACTGA ATTACGSTAT TTTCCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATTGCG CTGATTTGTG TGCGTATCGT ACAGAGTAAG
151 AAACAGTCGG AWAGCGGCAG TTGCGTACTG ACGGATTTTT CGGAAAATTA
201 TAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CACAACATCA GGAAAAAGAG GAAAAGAAAA AGGAAAAGCC GGAAGCCAAA
```

```

301 GCAGAGAAAA A.CGTTTGAA GGAGGGTGGG GAGAAATCTG CCGAAACGCA
351 nAAATCACGC CTTTTGTGT TGGANNNNNN NNNNNNNNNN NNNNNNNNNN
401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNGCGAGCGG CGGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGCTTCC GCTCCGTTTG CGATTGTCGG
651 TTCGGTGGGT GTGGTGGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAGAG GGCACACAGA AATTCCGACA
801 GGAACCTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTG AAAACAAACA GGTATTCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTCAGGC GGAAGCTTCT
1051 GTTGAAGAGT TGTTCGCAA ACTTGTC AAC CGGCGGGCGG ATGTGATGT A
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 22; ORF 005>:

```

m005.pep
1 MDNIDMFPE QEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQXSGSVVL TDFSENYKKQ RQSFEAFFLS GEEAQHQEKE EKKKEKAEAK
101 AEKXRLKEGG EKSAETXKSR LFVLXXXXXX XXXXXXXXXX XXXXXXXXXX
151 XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXXXXXXX XXXXASGGYM
201 MACVADKIAS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG
301 RQALALNLID EISTSDDLLL KAFENKQVIE VKYQEKQSLI QRIQLQAEAS
351 VEKLFKLVN RRADVM*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 23>:

```

a005.seq
1 ATGGACAATA TTGACATGTT CATGCCTGAA CAAGAGGAAA TCCAATCAAT
51 GTGGAAAGAA ATTTTACTGA ATTACGGTAT TTTCTGCTC GAACTGCTTA
101 CCGTGTTCGG CGCAATGCG CTGATTGTGT TGGCTATCGT ACAGAGTAAG
151 AAACAGTCGG AAAGCGGCAG TGTCGTACTG ACGGATTTTT CGGAAAATTA
201 TAAAAAACAG CGGCAATCGT TTGAAGCATT CTTTTTAAGC GGGGAAGAGG
251 CAAAACATCA GAAAAAGAG GAAAAGAAA AGGAAAAGGC GGAAGCCAAA
301 GCAGAGAAAA AGCGTTTGAA GGAGGGTGGG GAGAAATCTT CCGAAACGCA
351 AAAATCCCGC CTTTTGTGT TGGATTTTGA CGGCGATTG TATGCACACG
401 CCGTAGAATC CTTGCGTCAT GAGATTACGG CGGTGCTTTT GATTGCCAAG
451 CCTGAAGATG AGGTTCTGCT TAGATTGGAA AGTCCGGGCG CGGTGGTTCA
501 CCGTTACGGT TTGGCGGCTT CGCAGCTTAG GCGTTTGC GCACGCAATA
551 TTCCGCTGAC CGTCGCGCTC GATAAGGTGG CGGCGAGCGG TGGTTATATG
601 ATGGCGTGTG TGGCGGATAA AATTGTTTCC GCTCCGTTTG CGATTGTCGG
651 TTCGGTGGGT GTTGTAGCGG AAGTACCGAA TATCCACCGC CTGTTGAAAA
701 AACATGATAT TGATGTGGAT GTGATGACGG CGGGCGAATT TAAGCGCACG
751 GTTACTTTTA TGGGTGAAAA TACGGAAGAG GGCACACAGA AATTCCGACA
801 GGAACCTGGAG GAAACGCATC AGTTGTTCAA GCAGTTTGTC AGCGAGAACC
851 GCCCTCAATT GGATATTGAG GAAGTGGCAA CGGGCGAGCA TTGGTTCGGT
901 CGGCAGGCGT TGGCGTTGAA CTTGATTGAC GAGATTTCGA CCAGTGATGA
951 TTTGTTGTTG AAAGCGTTT AAAACAAACA GGTATTCGAA GTGAAATATC
1001 AGGAGAAGCA AAGCCTGATC CAGCGCATTG GTTTCAGGC GGAAGCTTCT
1051 GTTGAAGAGT TGTTCGCAA ACTTGTC AAC CGGCGGGCGG ATGTGATGTA
1101 G

```

This corresponds to the amino acid sequence <SEQ ID 24; ORF 005.a>:

```

a005.pep
1 MDNIDMFPE QEEIQSMWKE ILLNYGIFLL ELLTVFGAIA LIVLAIVQSK
51 KQSESGSVVL TDFSENYKKQ RQSFEAFFLS GEEAKHQEKE EKKKEKAEAK
101 AEKXRLKEGG EKSSSETQKSR LFVLDFDGD L YAHAVESLRH EITAVLLIAK
151 PEDEVLLRLE SPGGVVHGYG LAASQLRRLR ERNIPLTVAV DKVAASGGYM
201 MACVADKIVS APFAIVGSVG VVAEVPNIHR LLKKHDIDVD VMTAGEFKRT
251 VTFMGENTEK GKQKFRQELE ETHQLFKQFV SENRPQLDIE EVATGEHWFG

```

301 RQALALNLID EISTSDDLLL KAFENKQVIE VKYQEKQSLI QRIQLQAEAS
 351 VEKLFKLVN RRADV*

m005/a005 79.2% identity over a 366 aa overlap

m005.pep	10	20	30	40	50	60
	MDNIDMFMEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQXSXSGSVVL					
a005	MDNIDMFMEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSESGSVVL					
	10	20	30	40	50	60
m005.pep	70	80	90	100	110	120
	TDFSENYKKQRQSFEAFFLSGEEAQHQEKEKKKEKAEAKAEKKRLKEGGEKSAETXKSR					
a005	TDFSENYKKQRQSFEAFFLSGEEAQHQEKEKKKEKAEAKAEKKRLKEGGEKSSSETQKSR					
	70	80	90	100	110	120
m005.pep	130	140	150	160	170	180
	LFVLXX					
a005	LFVLDFDGLYAHAVESLRHEITAVLLIAKPEDEVLLRLESPPGGVVHGYGLAASQLRRLR					
	130	140	150	160	170	180
m005.pep	190	200	210	220	230	240
	XXXXXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLKKHDIDVD					
a005	ERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAIVGSVGVVAEVPNIHRLKKHDIDVD					
	190	200	210	220	230	240
m005.pep	250	260	270	280	290	300
	VMTAGEFKRTVTFMGENTEKKGKQKFRQELEETHQLFKQFVSENRPOLDIEEVATGEHWFG					
a005	VMTAGEFKRTVTFMGENTEKKGKQKFRQELEETHQLFKQFVSENRPOLDIEEVATGEHWFG					
	250	260	270	280	290	300
m005.pep	310	320	330	340	350	360
	RQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFKLVN					
a005	RQALALNLIDEISTSDDLLLKAFENKQVIEVKYQEKQSLIQRIGLQAEASVEKLFKLVN					
	310	320	330	340	350	360
m005.pep	RRADV					
a005	RRADV					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 005 shows 77.0% identity over a 366 aa overlap with a predicted ORF (ORF 005.ng) from *N. gonorrhoeae*:

m005/g005

m005.pep	10	20	30	40	50	
	MDNIDMFMEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQXSXSGSV					
g005	MGMDNIDMFMEQEEIQSMWKEILLNYGIFLLELLTVFGAIALIVLAIVQSKKQSESGSV					
	10	20	30	40	50	60
m005.pep	60	70	80	90	100	110
	VLTDSENYKKQRQSFEAFFLSGEEAQHQEKEKKKEKAEAKAEKKRLKEGGEKSAETXK					
g005	VLTDSENYKKQRQSFETFLSEEETKHQEKKEKKKEKAEAKAEKKRLKEGGEKSAETQK					
	70	80	90	100	110	120

```

m005.pep      120      130      140      150      160      170
                SRLFVLXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
                |||||
g005           130      140      150      160      170      180
                SRLFVLDFDGLYAHAVESLRHEITAVLLIAKPEDEVLLRLLESPGGVVHGYGLAASQLRR

m005.pep      180      190      200      210      220      230
                XXXXXXXXXXXXXASGGYMMACVADKIASAPFAIVGSVGVVAEVPNIHRLKKHDID
                : |||||:|||||:|||||:|||||:|||||
g005           190      200      210      220      230      240
                LRERNIPLTVAVDKVAASGGYMMACVADKIVSAPFAVIGSVGVVAEVPNIHRLKKHDID

m005.pep      240      250      260      270      280      290
                VDVMTAGFEKRTVTFMGENTEKKGKQKFRQEELEETHQLFKQFVSENRPQLDIEEVATGEHW
                |||||:|||||:|||||:|||||:|||||
g005           250      260      270      280      290      300
                VDVMTAGFEKRTVTFMGENTEKKGKQKFRQEELEETHQLFKQFVSENRPGLDIEKIATGEHW

m005.pep      300      310      320      330      340      350
                FGRQALALNLIDEISTSDDLLLKAFENQVIEVKYQEKQSLIQRIGLQAEASVEKLFACL
                |||||:|||||:|||||:|||||:|||||
g005           310      320      330      340      350      360
                FGRQALALNLIDEISTSDDLLLKAFENQVIEVKYQEKRSLIQRIQLQAEASVEKLFACL

m005.pep      360
                VNRRADVMX
                |||||
g005           VNRRADVMX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 25>:

```

g006.seq
1  ATGCTGCTGG TGCTggaatt ttggttCGGc gtGtCGGCGG TGGGCatact
51  tgCGTTGTTT TTATGGCttt TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCCG CCTGAACAAC AGCTTGGAAC gcgACAACCA CTTTATCCGA
151 AAAGGCGACG AGCGGCAGCT GTACCGCCAT TACGGACTGG TTCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCG
251 CGGCGATGGG TATTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACGGCA GCGCGGGGCA TATTTATTCG GTCGGCACTT ATCTGTGGAT
351 GTTTGCCATG AGTTTGACG ATGTGCCGCG ATTGGTCGAA CAATATCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTCGGAACG GAACATCAAA
451 GCCGGAACCTT GA

```

This corresponds to the amino acid sequence <SEQ ID 26; ORF 006.ng>:

```

g006.pep
1  MLLVLEFWFG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
51  KGDERQLYRH YGLVSRRLVL ISNREAFGYL CVGAAMGILF GFAFVMTLK
101 GYGSAghiys VGTYLWMFAM SLDDVPRLVE QYSLNKDIGQ RIEWSERNIK
151 AGT*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 27>:

```

m006.seq
1  ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTGCGCGG TGGGCATACT
51  TGCGTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCCG CCTGAACAAC AGCTTGGAAC GCGACAACCA CTTTATCCGA
151 AAAGGCGACC GCGGCAGCT GTACCGCCAT TACGGACTGC TTCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCGTCGGCA
251 CGGCGATGGG TATTTGTTC GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACAGCA GCGCGGGGCA TGTCTATTCTG GTCGGCACTT ATCTGTGGAT

```

```

351  GTTTGCCATG AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401  ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGGAACG GAACATCAAA
451  GCCGGAAGTTGA

```

This corresponds to the amino acid sequence <SEQ ID 28; ORF 006>:

```

m006.pep
1   MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLNN SLERDNHFIR
51  KGDRRQLYRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101 GYSSAGHVYS VGTYLWMFAM SLDDVPRLVE QYSLNKDIGQ RIEWSERNIK
151 AGT*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 29>:

```

a006.seq
1   ATGCTGCTGG TGCTGGAATT TTGGGTCGGC GTGTCGGCGG TGGGCATACT
51  TGCCTTGTTT TTATGGCTTT TGCCACGTTT TGCCGCCATC AGCGAAAACC
101 TGTATTTCGG CCTGAAGAAC AGCTTGAAC GCGACAACCA CTTTATCCGA
151 AAAGGCGACG AGCGGCAGCT GGACCGCCAT TACGGACTGC TTGCGCGCCT
201 GCGTGTGCTG ATTTCCAACC GCGAAGCCTT CGGCTATCTC TGCCTCGGCA
251 CGGCGATGGG TATTTTGTTT GGCTTTGCTT TTGTGATGAT GACGCTCAAA
301 GGCTACAGCA GCGCGGGGCA TGTCTATTCG GTCGGCACTT ATCTGTGGAT
351 GTTTGCCATA AGTTTGGACG ACGTGCCGCG ATTGGTCGAA CAATATTCCA
401 ATTTGAAAGA CATCGGACAA CGGATAGAGT GGTGGAACG GAACATCAAA
451 GCCGGAAGTT GA

```

This corresponds to the amino acid sequence <SEQ ID 30; ORF 006.a>:

```

a006.pep
1   MLLVLEFWVG VSAVGILALF LWLLPRFAAI SENLYFRLKN SLERDNHFIR
51  KGDERQLDRH YGLLARLRVL ISNREAFGYL CVGTAMGILF GFAFVMMTLK
101 GYSSAGHVYS VGTYLWMFAI SLDDVPRLVE QYSLNKDIGQ RIEWSERNIK
151 AGT*

```

m006/a006 96.7% identity over a 153 aa overlap

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVS					
a006	MLLVLEFWVGVS					
	10	20	30	40	50	60
m006.pep	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAM					
a006	YGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSAGHVYSVGTYLWMFAI					
	70	80	90	100	110	120
m006.pep	SLDDVPRLVEQYSLNKDIGQRIEWSERNIKAGTX					
a006	SLDDVPRLVEQYSLNKDIGQRIEWSERNIKAGTX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 006 shows 95.4% identity over a 153 aa overlap with a predicted ORF (ORF 006.ng) from *N. gonorrhoeae*:

m006/g006

	10	20	30	40	50	60
m006.pep	MLLVLEFWVGVS					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 31>:

This corresponds to the amino acid sequence <SEQ ID 32; ORF 006-1.ng>:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 33>:

```
m006-1.seq
1  ATGTGAAAAA  TGTGAAACA  CATAGCCCAA  ACCCACCGBA  AGCGATTGAT
51  TGGCATCATTT  TCCCTGGTCG  GACTGGAAAA  CCTTTTGATG  CTGGTGATCAT
101 CGGTGATTTGG  CGGCGGGGCG  ATCAATCCCG  TGATTGCGGG  GGAGGTGTGG
151 CAGGCGTTGCG  TGTACGCTTT  GGTGTGCTT  TTGATGTGGC  TGGTCCGTGC
201 GGTGCGGGCG  ATTGCCGATA  CGCGCAGCTT  TACGCGGATT  TATACCGAAA
251 TCCGCGTGCC  GGTCTGTTG  GAACAGCGCG  AGCGACAAGT  CCGCATTCG
301 GCGGTAACCTG  CGCGGGTTGC  CCTGTCGCGT  GAGTTTGTC  GCTTTTTTGA
351 AGAACACCTG  CCGATTGCCG  CGACATCCGT  CGTATCCATA  TTCGCGCGCT
401 GCATCATGCT  GCTGTGCTG  GAATTTGGG  TCGGCGGTGC  GCGCGTGGCG
451 ATACTTGCGT  TGTTTTATG  GCTTTTGCCA  CGTTTGGCCG  CCGATCAGCA
501 AAACCTGTAT  TTCGCGCTGA  ACAACAGCTT  GGAACCGGAC  AACTCACTTA
551 TCCGAAAAGG  CGACGGCGG  CAGCTGTACC  GCCATTACGG  ATCTGCTGCG
601 CGCCTGCGTG  TGCTGATTT  CAACCGCGAA  GCCTTCGCGT  ATCTCTGCGT
```

```

651 CGGCACGGCG ATGGGTATTT TGTTCGGCTT TGCTTTTGTG ATGATGACGC
701 TCAAAGGCTA CAGCAGCGCG GGGCATGTCT ATTCGGTCGG CACTTATCTG
751 TGGATGTTTG CCATGAGTTT GGACGACGTG CCGCGATTGG TCGAACAATA
801 TTCCAATTTG AAAGACATCG GACAACGGAT AGAGTGGTCG GAACGGAACA
851 TCAAAGCCGG AACTTGA

```

This corresponds to the amino acid sequence <SEQ ID 34; ORF 006-1>:

m006-1.pep

```

1 MWKMLKHIAQ THRKLIGTF SLVGLENLLM LVYPVFGGRA INAVIAGEVW
51 QALLYALVVL LMWLVGAVRR IADTRTFTRI YTEIAVPVVL EQRQRQVPHS
101 AVTARVALSR EFVSFFEEHL PIAATSVVSI FGACIMLLVL EFWVGVSAVG
151 ILALFLWLLP RFAAISENLY FRLNNSLERD NHFIRKGDRL QLYRHYGLLA
201 RLRVLISNRE AFGYLCVGTA MGILFGFAFV MMTLKGYS SA GHVYSVGTYL
251 WMFAMSLDDV PRLVEQYSNL KDIGQRIEWS ERNIKAGT*

```

m006-1/g006-1 95.5% identity in 288 aa overlap

m006-1.pep	10	20	30	40	50	60
	MWKMLKHIAQTHRKLIGTFSLVGLENLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL					
g006-1	MWKMLKHIAKTHRKLIGTFSPVGLENLLMLGYPVFGGWAINAVIAGRVWQALLYALVVF					
	10	20	30	40	50	60
m006-1.pep	70	80	90	100	110	120
	LMWLVGAVRRRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL					
g006-1	LMWLVGAAARRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL					
	70	80	90	100	110	120
m006-1.pep	130	140	150	160	170	180
	PIAATSVVSIFGACIMLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERD					
g006-1	PIAATSVVSIFGACIMLLVLEFWVGVSAVGILALFLWLLPRFAAISENLYFRLNNSLERD					
	130	140	150	160	170	180
m006-1.pep	190	200	210	220	230	240
	NHFIRKGDRLQLYRHYGLLARLVLISNREAFGYLCVGTAAMGILFGFAFVMMTLKGYS SA					
g006-1	NHFIRKGDRLQLYRHYGLVSRRLVLISNREAFGYLCVGTAAMGILFGFAFVMMTLKGYS SA					
	190	200	210	220	230	240
m006-1.pep	250	260	270	280	289	
	GHVYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX					
g006-1	GHIYSVGTYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX					
	250	260	270	280		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 35>:

a006-1.seq (partial)

```

1 ..AGCCAAAACC ACCGCAAGCG ATTGATTGGC ACATTTTTTC TGGTCGGACT
51 GGAAAACCTT TTGATGCTGG TGTATCCGGT GTTTGGCGGC TGGGCGATTA
101 ATGCCGTGAT TGCAGGGCAG GCGTGGCAGG CGTTGCTGTA CGCTTTGGTT
151 GTGCTTTTGA TGTGGCTGGT CCGTGCGGCG CGGCGGATTG CCGATACGCG
201 CACGTTTACG CGGATTTATA CCGAAATCGC CGTGCCGTT GTGTTGGAAC
251 AGCGGCAGCG GCAAGTCCCG CATTGCGCGG TAACTGCGCG GGTGCGCTG
301 TCGCGTGAGT TTGTCAGCTT TTTGAAGAA CACCTGCCGA TTGCCGCGAC
351 ATCCGTCGTA TCCATATTCG GCGCGTGCAT CATGCTGCTG GTGCTGGAAT
401 TTTGGGTCGG CGTGTGCGCG GTGGGCATAC TTGCGTTGTT TTTATGGCTT
451 TTGCCACGTT TTGCCGCCAT CAGCGAAAAC CTGTATTTC GCCTGAAGAA
501 CAGCTTGGAA CGCGACAACC ACTTTATCCG AAAAGGCGAC GAGCGGCAGC
551 TGGACCGCCA TTACGGACTG CTGCGCGGCC TCGGTGTGCT GATTTCCAAC
601 CGCGAAGCCT TCGGCTATCT CTGCGTCGGC ACGGCGATGG GTATTTGTGT
651 CGGCTTTGCT TTTGTGATGA TGACGCTCAA AGGCTACAGC AGCGCGGGGC

```

701 ATGTCTATTC GGTCCGCACT TATCTGTGGA TGTTCGCCAT AAGTTTGGAC
 751 GACGTGCCGC GATTGGTCGA ACAATATTCC AATTTGAAAG ACATCGGACA
 801 ACGGATAGAG TGGTCGAAAC GGAACATCAA AGCCGGAAC TGA

This corresponds to the amino acid sequence <SEQ ID 36; ORF 006-1.a>:

a006-1.pep (partial)
 1 ..SQNHRKRLIG TFFLVGLENL LMLVYPVFGG WAINAVIAGO AWOALLYALV
 51 VLLMWLVGAA RRIADTRTFT RIYTEIAVPV VLEQRQRQVP HSAVTARVAL
 101 SREFVSFFEE HLPAAATSVV SIFGACIMLL VLEFWVGUSA VGILALFLWL
 151 LPRFAAISEN LYFRLKNSLE RDNHFIRKGD ERQLDRHYGL LARLRVLISN
 201 REAFGYLCVG TAMGILEGFA FVMMTLKGYS SAGHVYSVGT YLWMFAISLD
 251 DVPRLVEQYS NLKDIGQRIE WSKRNIKAGT *

a006-1/m006-1 95.7% identity in 280 aa overlap

	10	20	30	40	50
a006-1.pep	SQNHRKRLIGTFFLVGLENLLMLVYPVFGGWAINAVIAGOAWOALLYALVVL				
	: : :				
m006-1	MWKMLKHIAQTHRKRLIGTFFLVGLENLLMLVYPVFGGRAINAVIAGEVWQALLYALVVL				
	10	20	30	40	50
	60	70	80	90	100
a006-1.pep	LMWLVGAAARRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL				
	:				
m006-1	LMWLVGAVRRIADTRTFTRIYTEIAVPVLEQRQRQVPHSAVTARVALSREFVSFFEEHL				
	70	80	90	100	110
	120	130	140	150	160
a006-1.pep	PIAATSVVSIFGACIMLLVLEFWVGVSAGVILALFLWLLPRFAAISENLYFRLKNSLERD				
m006-1	PIAATSVVSIFGACIMLLVLEFWVGVSAGVILALFLWLLPRFAAISENLYFRLNNSLERD				
	130	140	150	160	170
	180	190	200	210	220
a006-1.pep	NHFIRKGDERQLDRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSA				
	:				
m006-1	NHFIRKGDERRQLYRHYGLLARLRVLISNREAFGYLCVGTAMGILFGFAFVMMTLKGYSSA				
	190	200	210	220	230
	240	250	260	270	280
a006-1.pep	GHVYSVGTLYLWMFAISLDDVPRLVEQYSNLKDIGQRIEWSKRNIKAGTX				
	:				
m006-1	GHVYSVGTLYLWMFAMSLDDVPRLVEQYSNLKDIGQRIEWSERNIKAGTX				
	250	260	270	280	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 37>:

g007.seq
 1 atgaACACAA CCGGACTGCC GACCGCCTTC ATCTTGCTCT GCCTCTGcgC
 51 CGCcGCTTCT GCCGccgaca acAGCatcat gaCaAAAGGG CAAAAAGTGT
 101 ACGAATCcAa ctGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
 151 ACTGCGtTTC CTccgctTTT CCggtcgGac tgtattatga acaAACCGCa
 201 cgTCCTgtctg cacagcatgg tcaaaggcAt cgacgggaca ttcaaagtgg
 251 agcggcaaaa cctacgacgg atttatgCcc gcaaccgcca tcagcgATGC
 301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

This corresponds to the amino acid sequence <SEQ ID 38; ORF 007.ng>:

g007.pep
 1 MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVEYNCA CHGKKGEGRG
 51 TAFPPLFRSD CIMNKPVLH HSMVKGIDGT FKVERQNLRR IYARNRHQRC
 101 GHCRRRHLYH ERL*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 39>:

m007.seq

```

1 ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCTTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACC GCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251 AACGGCAAAA CCTACAACGG ATTCATGCCC GCAACGCCA TCAGCGATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 40; ORF 007>:

m007.pep

```

1 MNTTRLPTAL VLGCFCAAA AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARNRHQRC
101 GCHRRRHLYH ERL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 41>:

a007.seq

```

1 ATGAACACAA CCCGACTGCC GACCGCCCTC GTCTTGGGCT GCCTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACC GCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTC.
251 AACGGCAAAA CCTACAACGG ATTCATGCCC GCCACTGCCA TCAGCGATGC
301 GGACATTGCC GCCGTCGCCA CTTATATCAT GAACGCCTTT GA

```

This corresponds to the amino acid sequence <SEQ ID 42; ORF 007.a>:

a007.pep

```

1 MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVXRQNLQR IHARHCHQRC
101 GCHRRRHLYH ERL*

```

m007/a007 97.3% identity over a 113 aa overlap

	10	20	30	40	50	60
m007.pep	MNTTRLPTALVLGCFCAAA AADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	: : : : : : : : : :					
a007	MNTTRLPTALVLGCLCAA AADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	10	20	30	40	50	60
	70	80	90	100	110	
m007.pep	FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGCHRRRHLYHERLX					
	: : : : : : : : :					
a007	FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARHCHQRCGCHRRRHLYHERLX					
	70	80	90	100	110	

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 007 shows 86.7% identity over a 113 aa overlap with a predicted ORF (ORF 007.ng) from *N. gonorrhoeae*:

m007/g007

	10	20	30	40	50	60
m007.pep	MNTTRLPTALVLGCFCAAA AADNSIMTKGQKVYESNCVACHGKKGEGRGTMFPPLYRSD					
	: : : : : : : : :					
g007	MNTTRLPTAFILCCLCAA AADNSIMTKGQKVYESNCIACHGKKGEGRGTAFPPFLFRSD					
	10	20	30	40	50	60
	70	80	90	100	110	
m007.pep	FIMKKPQVLLHSMVKGINGTIKVXRQNLQRIHARNRHQRCGCHRRRHLYHERLX					
	: : : : : : : : :					

g007 CIMNKPHVLLHSMVKGIDGTFKVERQNLRRYARNRHQRCGHCRRRHLYHERL
 70 80 90 100 110

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 43>:

g007-1.seq (partial)

```

1  ATGAACACAA CCCGACTGCC GACCGCCTTC ATCTTGTGCT GCCTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCATCGCC TGCCACGGCA AGAAAGGGGA AGGGCGCGGC
151 ACTGCGTTTC CTCCGCTTTT CCGGTCGGAC TATATTATGA ACAAACCGCA
201 CGTCCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAG ACGTAAACA GGCAAAAGGC AAAAAAAC.
```

This corresponds to the amino acid sequence <SEQ ID 44; ORF 007-1.ng>:

g007-1.pep (partial)

```

1  MNTTRLPTAF ILCCLCAAAS AADNSIMTKG QKVEYENCIA CHGKKGEGRG
51  TAFPPLFRSD YIMNKPHVLL HSMVKINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKG KKN...
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 45>:

m007-1.seq

```

1  ATGAACACAA CCCGACTGCC GACCGCCTTC GTCTTGGGCT GCTTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAAACGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CAACCGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAG ACGTAAACA GGCAAAAGC AAAAAAACT
401 AA
```

This corresponds to the amino acid sequence <SEQ ID 46; ORF 007-1>

m007-1.pep

```

1  MNTTRLPTAL VLGCFCAAAS AADNSIMTKG QKVEYENCAV CHGKKGEGRG
51  TMFPPLYRSD FIMKKPVLL HSMVKINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKS KKN*
```

m007-1 / g007-1 91.7% identity in 133 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGCFCAAAS	AADNSIMTKGQKVEYENCAV	CHGKKGEGRGTMFPPLYRSD			
g007-1	: : : : : : :	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPVLLHSMVKINGT	IKVNGKTYNGFMPATAISD	ADIAAVATYIMNAFDNGGGSV			
g007-1	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :	: : : : : : : : :
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKSKKNX					
g007-1	: : : : : : :					
	TEKDVQAKGKKN					
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 47>:

a007-1.seq (partial)

```

1  ATGAACACAA CCCGACTGCC GACCGCCTTC GTCTTGGGCT GCCTCTGCGC
51 CGCCGCTTCT GCCGCCGACA ACAGCATCAT GACAAAAGGG CAAAAAGTGT
101 ACGAATCCAA CTGCGTCGCC TGCCACGGCA AAAAGGGCGA AGGCCGCGGA
```

```

151 ACCATGTTTC CGCCGCTCTA CCGCTCCGAC TTCATCATGA AAAAACCGCA
201 GGTGCTGCTG CACAGCATGG TCAAAGGCAT CAACGGTACA ATCAAAGTCA
251 ACGGCAAAAC CTACAACGGA TTCATGCCCG CCACTGCCAT CAGCGATGCG
301 GACATTGCCG CCGTCGCCAC TTATATCATG AACGCCTTTG ACAACGGCGG
351 CGGAAGCGTT ACCGAAAAAG ACGTAAAACA GGCAAAAAAC AAAAAA..

```

This corresponds to the amino acid sequence <SEQ ID 48; ORF 007-1.a>:

a007-1.pep (partial)

```

1 MNTTRLPTAL VLGCLCAAAS AADNSIMTKG QKVYESNCVA CHGKKGEGRG
51 TMFPPLYRSD FIMKKPQVLL HSMVKGINGT IKVNGKTYNG FMPATAISDA
101 DIAAVATYIM NAFDNGGGSV TEKDVQAKN KK..

```

m007-1/a007-1 98.5% identity in 132 aa overlap

	10	20	30	40	50	60
m007-1.pep	MNTTRLPTALVLGCLCAAAS	AAADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
a007-1	MNTTRLPTALVLGCLCAAAS	AAADNSIMTKGQKVYESNCVACHGKKGEGRG	TMFPPLYRSD			
	10	20	30	40	50	60
	70	80	90	100	110	120
m007-1.pep	FIMKKPQVLLHSMVKGINGTIKVN	GKTYNGFMPATAISDADIAAVATYIM	NAFDNGGGSV			
a007-1	FIMKKPQVLLHSMVKGINGTIKVN	GKTYNGFMPATAISDADIAAVATYIM	NAFDNGGGSV			
	70	80	90	100	110	120
	130					
m007-1.pep	TEKDVQAKSKKNX					
a007-1	TEKDVQAKNKK					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 49>:

g008.seq

```

1 ATGAACAACA GACATTTTGC CGTCATcgCC TTGGGCAGCA ACCTTGACAA
51 CCCCGCACAA CAAATacgcg gcgcattaga cgcgctctcg tcccatcctg
101 acatccggct tgaacaggtt tcctcactgt aTatgaccgc acctgtcggg
151 tacgAcaaTC agcccgATTT CATCaatgcc gTCTgcaccg TTTCCACCAC
201 CtTGGACGGC ATTGcccTGC TTGCCgaACT CAAccgTATC GAAGCCGATT
251 TCGGACGCGA aCGCAGTTTC CGCAATGCAC CGCGCACATT GGATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGCC TTACCCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATACGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAAATACG GAAAGGTTGT CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGACA GTTAA

```

This corresponds to the amino acid sequence <SEQ ID 50; ORF 008.ng>:

g008.pep

```

1 MNNRHFAVIA LGSNLDNPAQ QIRGALDALS SHPDIRLEQV SSLYMTAPVG
51 YDNQPDFINA VCTVSTTLDG IALLAELNRI EADFRGRERSF RNAPRTLDD
101 IIDFDGISSD DPRLTLPHPR AHERSEVIRP LAEILPDFIL GKYGKVVELS
151 KRLGNQGI RL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 51>:

m008.seq

```

1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
51 CCCTGCTCAA CAGGTACGCG CCGCATTTGA CACGCTGTCG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTT TGTCATGCC GTCTGCACCG TTTCACCAC
201 TCTGGACGGC ATTGCCyTGC TTGCCGAAct CAACCGTATC GAGGCTGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GkATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACACsCGAC TcACctTGCC

```

```

351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATCCGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTGTTTTA GGAAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGyTGG GCAATCAAGG TATCCGTCTT TTACCGGACA GGTAATT

```

This corresponds to the amino acid sequence <SEQ ID 52; ORF 008>:

```

m008.pep
  1 MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSlyMTAPVG
 51 YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADfGRERSF RNAPRTLxLD
101 IIDFDGISSD DTRLTLPHFR AHERSFVIRP LAEILPDFVL GKHGKVAELS
151 KRLGNQGIRL LPDR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 53>:

```

a008.seq
  1 ATGAACAACA GACATTTTGC CGTCATCGCC CTGGGCAGTA ATCTTGAAAA
 51 CCCTGCCCAA CAGGTACGCG CCGCATTGGA CACGCTGTCG TCCCATCCTG
101 ACATCCGTCT TAAACAGGCT TCCTCACTGT ATATGACCGC GCCCGTCGGT
151 TACGACAATC AGCCCGATTG CGTCAATGCC GTCTGCACCG TTTCCACCAC
201 CTTGGACGGC ATTGCCCTGC TTGCCGAAC CAACCGTATC GAAGCCGATT
251 TCGGACGCGA ACGCAGCTTC CGCAACGCGC CGCGCACATT GGATTGGAC
301 ATTATCGACT TTGACGGCAT CTCCAGCGAC GACCCCGGAC TCACCCTGCC
351 GCATCCGCGC GCGCACGAAC GCAGTTTCGT CATAcGCCCT TTGGCAGAAA
401 TCCTCCCTGA TTTTATTTTG GGAAAACACG GAAAGGTTGC CGAATTGTCA
451 AAACGGCTGG GCAATCAAGG CATCCGTCTT TTACCGGATA AGTAA

```

This corresponds to the amino acid sequence <SEQ ID 54; ORF 008.a>:

```

a008.pep
  1 MNNRHFAVIA LGSNLENPAQ QVRAALDTLS SHPDIRLKQA SSlyMTAPVG
 51 YDNQPDFVNA VCTVSTTLDG IALLAELNRI EADfGRERSF RNAPRTLdLD
101 IIDFDGISSD DRLTLPHFR AHERSFVIRP LAEILPDFIL GKHGKVAELS
151 KRLGNQGIRL LpDK*

```

m008/a008 97.6% identity over a 164 aa overlap

	10	20	30	40	50	60
m008.pep	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
a008	MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA					
	10	20	30	40	50	60
m008.pep	70	80	90	100	110	120
	VCTVSTTLDGIALLAELNRIEADfGRERSFRNAPRTLxLDIIDFDGISSDDTRLTLPHFR					
a008	VCTVSTTLDGIALLAELNRIEADfGRERSFRNAPRTLdLDIIDFDGISSDDPRLTLPHFR					
	70	80	90	100	110	120
m008.pep	130	140	150	160		
	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIRLLPDRX					
a008	AHERSFVIRPLAEILPDFILGKHGKVAELSKRLGNQGIRLLPDKX					
	130	140	150	160		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 008 shows 92.7% identity over a 164 aa overlap with a predicted ORF (ORF008.ng) from *N. gonorrhoeae*:

```

m008/g008
      10      20      30      40      50      60
m008.pep  MNNRHFAVIALGSNLENPAQQVRAALDTLSSHPDIRLKQASSLYMTAPVGYDNQPDFVNA
g008      MNNRHFAVIALGSNLDNPAQQIRGALDALSSHPDIRLEQVSSLYMTAPVGYDNQPDFINA

```

183

	10	20	30	40	50	60
	70	80	90	100	110	120
m008.pep	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDTRLTLPHPR					
g008	VCTVSTTLDGIALLAELNRIEADFGRRERSFRNAPRTLXLDIIDFDGISSDDPRLTLPHPR					
	70	80	90	100	110	120
	130	140	150	160		
m008.pep	AHERSFVIRPLAEILPDFVLGKHGKVAELSKRLGNQGIIRLLPDRX					
g008	AHERSFVIRPLAEILPDFILGKYGKVVELSKRLGNQGIIRLLPDRX					
	130	140	150	160		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 55>:

m009.seq
 1 ATGCCCGCG CTGCCGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
 51 CGAACAAAAT ACCCATCGCC GCGCCGACGC AGAGATAGCC GAAGGCTTCG
 101 CGGTTGGAAA TCAGCACACG CAGGCGCGAA ACCAGTCCGT AATGGCGGTA
 151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTcg cGTTCCAAGC
 201 TGTTGTTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
 251 AaaaGCCATA A

This corresponds to the amino acid sequence <SEQ ID 56; ORF 009.ng>:

g009.pep
 1 MPRAAFAFER HHHKSKEQN THRRADAEIA EGFVGNQHT QARNQSVMAV
 51 QLPLVAFSDK VVVAFAQVVQ AEIQVFADGG KTWQKP*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 57>:

m009.seq
 1 ATGCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
 51 CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
 101 CGGTTGGAAA TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTA
 151 CAGCTGCCGC CGGTCGCCTT TTCGGATAAA GTGGTTGTGC CGTTCCAAGC
 201 TGTTGTTTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
 251 AAAAGCCATA A

This corresponds to the amino acid sequence <SEQ ID 58; ORF 009>:

m009.pep
 1 MPRAAFAFER HHHKSKEQN THRRADAEIA EGFVGNQHT QARKQSVMAV
 51 QLPPVAFSDK VVVAFAQVVQ AEIQVFADGG KTWQKP*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 009 shows 97.7% identity over a 86 aa overlap with a predicted ORF (ORF 009.ng) from *N. gonorrhoeae*:

m009/g009

	10	20	30	40	50	60
m009.pep	MPRAAFAFERHHHKSKEQNTHRRADAEIAEGFVGNQHTQARKQSVMAVQLPPVAFSDK					
g009	MPRAAFAFERHHHKSKEQNTHRRADAEIAEGFVGNQHTQARNQSVMAVQLPLVAFSDK					
	10	20	30	40	50	60
	70	80				
m009.pep	VVVAFAQVVQAEIQVFADGGKTWQKXPX					
g009	VVVAFAQVVQAEIQVFADGGKTWQKXPX					
	70	80				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 59>:

```
a009.seq
1  ATGCCCCGCG CTGCTGTAGC CTTTGAGCGT CATCATCACA AAAGCAAAGC
51  CGAACAAAAT ACCCATCGCC GTGCCGACGC AGAGATAGCC GAAGGCTTCG
101 CGGTTGGAAT TCAGCACACG CAGGCGCGCA AGCAGTCCGT AATGGCGGTC
151 CAGCTGCCGC TCGTCGCCTT TTCGGATAAA GTGGTTGTCG CGTTCCAAGC
201 TGTTCCTCAG GCGGAAATAC AGGTTTTCGC TGATGGCGGC AAAACGTGGC
251 AAAAGCCATA A
```

This corresponds to the amino acid sequence <SEQ ID 60; ORF 009.a>:

```
a009.pep
1  MPRAAVAFAER HHHKSKAEQN THRRADAEIA EGFVGNQHT QARKQSVMAV
51  QLPLVAFSDK VVVAFAQVLO AEIQVFADGG KTWQKP*
```

m009/a009 97.7% identity over a 86 aa overlap

	10	20	30	40	50	60
m009.pep	MPRAAVAFAERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPPVAFSDK					
a009	MPRAAVAFAERHHHKSKAEQNTHRRADAEIAEGFAVGNQHTQARKQSVMAVQLPLVAFSDK					
	10	20	30	40	50	60
	70	80				
m009.pep	VVVAFAQVAVQAEIQVFADGGKWTQKPKX					
a009	VVVAFAQVLOAEIQVFADGGKWTQKPKX					
	70	80				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 61>:

```
g010.seq
1  ATGGGTTTTC CTGTTGCGAA GTTTGATGCC GTGATTGTCG GCGGTGGCGG
51  TGCAAGTTTA CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTGTGC TAAAGTGTTC CCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GCGGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCCAAA TTTATCAGCG
351 TCCTTTTCGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTG TATTTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTTGCGC CCGTGCGGGC ATTCCGTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGTATTTC TGTTGAacgc
801 cgacggcgaa cgcTTTATGG AAcgctatgc GCcgACCGta aAagaCTTG
851 CTTCTCGCga cgtGGTTTCA CgcgcGatgG CGatggaAAt ctatgaaggt
901 cgcggctgTG GtaaaAAaCaA agaCCacgtC TACTGAAAA TCGACcAtAt
951 cggTGcAGAA AAAATTATGG AAAAATGCC GGGCATCCGC GAGATTCCA
1001 TTCagtttgc cGGTATCGAT CCGATTAAAG ACCCGATTcc ggttgTGCCG
1051 ACTACCCACT ATATGATGGG CCGCATTCCg aCCAATTATC ACGGTGAAGT
1101 TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT
1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTC cgcccaaccc cccggtga
```

This corresponds to the amino acid sequence <SEQ ID 62; ORF 010.ng>:

g010.pep

```

1  MGFVPRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDWHMY DTVKGSWDLG DQDAIEFMCR AAPEAVIELE
101 HMGMFPDRVE SGKIYQRPFG GHAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRAQTQF FVEWTAQDLI RDENGDVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVUGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RCGGKNKDHV LLKIDHIGAE KIMEKLPGR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVVP QGDEYEVFVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTFR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 63>:

```

m010.seq (PARTIAL)
1  ..nTCCAATTAT CCAAATCCGG TCTGAATTGT GCCGTTTTGT CTAAAGTGT
51  CCCGACCCGT TCGCATACCG TAGCGGCGCA GGGCGGTATT TCCGCCTCTn
101 TGGGTAATGT GCAGGAAGAC CGTGGGACT GGCACATGTA CGATACCGTG
151 AAAGGTTCCG ACTGTTGGG CGACCAAGAT GCGATTGAGT TTATGTGCCG
201 CGCCGCGCCT GAAGCCGTAA TTGAGTTGGA ACACATGGGT ATGCCTTTTG
251 ACCGTGTGGA AAGCGGTAAA ATTTATCAGC GTCCTTTCGG CGGCCATACT
301 GCCGAACACG GTAAACGCGC GGTAGAACGC GyCTGTGCGG TTGCCGACCG
351 TACAGGTCAT GCGATGCTGC ATACTTTGTA CCAACAAAAC GTCGTGCGCA
401 ATACGCAATT CTTGTGGAA TGGACGGCAC AAGATTTGAT TCGTGATGAA
451 AACGGCGATG TCGTCGGCGT AACCGCCATG GAAATGGAAA CCGGCGAAGT
501 TTATATTTTC CACGCTAAAG CTGTGATGTT TGCTACCGGC GCGGCGGTTC
551 GTATTTATGC GTCTTCTACC AATGCCTATA TGAATACCGG CGATGGTTTG
601 GGTATTTGTG CGCGTGCAGG TATCCCGTTG GAAGACATGG AATTCTGGCA
651 ATTCCAGCCG ACCGGCGTGG CGGGTGCGGG CGTGTGATT ACCGAA...

```

This corresponds to the amino acid sequence <SEQ ID 64; ORF 010>:

```

m010.pep (PARTIAL)
1  ..XQLSKSGLNC AVLSKVFPR SHTVAAQGGI SASXGNVQED RWDWHMYDTV
51  KGSWDLGDQD AIEFMCRAP EAVIELEHMG MPFDRVESGK IYQRPFGGHT
101 AEHGKRAVER XCAVADRTGH AMLHTLYQON VRANTQFFVE WTAQDLIRDE
151 NGDVVGVTAM EMETGEVYIF HAKAVMFATG GGGRIYASST NAYMNTGDGL
201 GICARAGIPL EDMFVQFQP TGVAGAGVLI TE...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 65>:

```

a010.seq
1  ATGGGCTTTC CTGTCGCAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GCGGATATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCG GCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACCGG
401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
451 CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGCGG TATTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGGTATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCATGG CGATGGAAT CTACGAAGGT
901 CCGGCTGCG GTAAAAACAA AGACCATGTC TTTACTGAAA TCGACCATAT
951 CCGCGCAGAA AAAATTATGG AAAAATGCC GGCATCCGC GAGATTTCCT
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCGG
1051 ACTACCCACT ATATGATGGG CGTATTCCG ACCAACTACC ATGGCGAAGT
1101 TGTCGTTTCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAACCT TTGCCTGCTA

```

```

1301 ATGCCGCGCA ACTGACCCGC CAACGTATCG AGCGTTTGA CAATCAAAC
1351 GATGGTGAAA ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401 ATTGCACGCC GCGGTGTTC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GGCGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTAATTGAA GTGGCGAAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATGCCAATA CCTTGTCTTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 66; ORF 010.a>:

```

a010.pep
1  MGFPVRKFDA VIVGGGGAGL RAXLQLSKSG LNCVLSKV F PTRSHTVAAQ
51  GGISASLGNV QEDRWDMHY DTVKGS DWLG DQDAIEFMCR AAPEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNV RANTQF FVEWTAQDLI RDENG DVVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLSRDVVS RAMAMEIYEG
301 RCGCKNKDHF LLKIDHIGAE KIMEKLP GIR EISIQFAGID PIKDPVPV
351 TTHYMMGGIP TNYHGEVVVP QGDEYEV PVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMIAA ERVKRTEIKD
501 KSKVWNTARI EALELDN LIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKR VY*

```

m010/a010 98.7% identity over a 231 aa overlap

```

m010.pep
10 20 30
XQLSKSGLNCVLSKVFPTRSHTVAAQGGISASXGNV
|||||
a010 MGFPVRKFDAVIVGGGGAGLRAXLQLSKSGLNCVLSKVFPTRSHTVAAQGGISASLGNV
10 20 30 40 50 60
40 50 60 70 80 90
m010.pep QEDRWDMHYDTVKGSDWLGDDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
|||||
a010 QEDRWDMHYDTVKGSDWLGDDQDAIEFMCRAAPEAVIELEHMGMPFDRVESGKIYQRPFG
70 80 90 100 110 120
100 110 120 130 140 150
m010.pep GHTAEHGKRAVERXC AVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVGV
|||||
a010 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNV RANTQFFVEWTAQDLIRDENG DVVGV
130 140 150 160 170 180
160 170 180 190 200 210
m010.pep TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
|||||
a010 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ
190 200 210 220 230 240
220 230
m010.pep FQPTGVAGAGVLITE
|:|||||
a010 FHPTGVAGAGVLITEGV RGE GILLNADGERFMERYAPT VKDLSRDVVS RAMAMEIYEG
250 260 270 280 290 300

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 010 shows 98.7% identity over a 231 aa overlap with a predicted ORF (ORF 010.ng) from *N. gonorrhoeae*:

m010.pep/g010.pep

m010.pep				10	20	30	
				XQLSKSGLNCAVLSKVFPTRSH	TVA	AQGGISASXGNV	
g010	MGFPVRKFD	AVIVGGGAGLRAAL	QLSKSG	LNC	AVLSKVFPTRSH	TVA	AQGGISASLGNV
	10	20	30	40	50	60	
m010.pep	40	50	60	70	80	90	
	QEDRWDWHMYDTVKGS	DWLGDQDAIEFMCRAA	PEAVIE	LEHMGMPFDR	VESGKIYQRPFG		
g010	QEDRWDWHMYDTVKGS	DWLGDQDAIEFMCRAA	PEAVIE	LEHMGMPFDR	VESGKIYQRPFG		
	70	80	90	100	110	120	
m010.pep	100	110	120	130	140	150	
	GHTAEHGKRAVERXCA	VA	DRTGHAMLHTLYQ	QNV	RANTQFFVEW	TAQDLIRD	ENG
g010	GHTAEHGKRAVERAC	VA	DRTGHAMLHTLYQ	QNV	RANTQFFVEW	TAQDLIRD	ENG
	130	140	150	160	170	180	
m010.pep	160	170	180	190	200	210	
	TAMEMETGEVYIFHAK	AVMFATGGGGRIYAS	STNAYMNTGDGLG	ICARAGI	PLEDM	EFWQ	
g010	TAMEMETGEVYIFHAK	AVMFATGGGGRIYAS	STNAYMNTGDGLG	ICARAGI	PLEDM	EFWQ	
	190	200	210	220	230	240	
m010.pep	220	230					
	FQPTGVAGAGVLITE						
	:						
g010	FHPTGVAGAGVLITE	GEVGRGEGGILLNAD	GERFMERYAPT	VDLASRD	VVSRAMAMEI	YEG	
	250	260	270	280	290	300	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 67>:

g010-1.seq..

```

1  ATGGGTTTTC CTGTTTCGCAA GTTTGATGCC GTGATTGTGCG GCGGTGGCGG
51  TGCAGGTTTA CGTGCAGCCC TCCAATTATC CAAATCCGGT TTGAATTGTG
101 CCGTTTGTGC TAAAGTGTTT CCGACCCGCT CGCATACCGT AGCGGCGCAG
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAGGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGCTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGTCGC GCTGCGCCTG AAGCGGTGAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGCGTTGAA AGCGGCAAAA TTTATCAGCG
351 TCCTTTCGGC GGACATACTG CCGAACATGG TAAACGTGCG GTAGAACGTG
401 CATGTGCGGT TGCCGACCGT ACCGGTCATG CGATGTTGCA TACTTTGTAC
451 CAACAAAACG TCCGTGCCAA TACACAATTC TTTGTGGAAT GGACGGCGCA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC GGGCGAAGTT TATATTTTCC ACGCCAAGGC CGTGATGTTT
601 GCTACCGGTG GCGGCGGTCT TATTATGCT TCTTCTACCA ATGCTTATAT
651 GAATACCGGT GACGGTTTGG GCATTGCGC CCGTGCGGGC ATTCCGTTGG
701 AAGATATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGCG GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGCTATTC TGTGAACGC
801 CGACGGCGAA CGCTTATAGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTGGTTTCA CGCGCATGG CGATGGAAT CTATGAAGGT
901 CGCGGCTGTG GTAAAAACAA AGACCACGTC TTACTGAAAA TCGACCATAT
951 CGGTGCAGAA AAAATTATGG AAAAAGTACC GGGCATCCGC GAGATTTCCT
1001 TTCAGTTTGC CGGTATCGAT CCGATTAAAG ACCCGATTCC GGTGTGCGCG
1051 ACTACCACT ATATGATGGG CGGCATTCCG ACCAATTATC ACGGTGAAGT
1101 TGTTGTTCCG CAAGGCGACG AGTACGAAGT ACCTGTAAAA GGCCTGTATG
1151 CCGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGTTTGGGT

```

1201 ACGAACTCCC TGCTGGACTT GGTGGTGTTT cgcccaaccc cccggtga

This corresponds to the amino acid sequence <SEQ ID 68; ORF 010-1.ng>:

g010-1.pep

```

1  MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVFP TRSHTVAAQ
51  GGISASLGNV QEDRWDRWHY DTVKGSDDLWQ DQDAIEFMCR APEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRRANTQF FVEWTAQDLI RDENGDDVGV TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE REMERYAPT V KDASRDVVS RAMAMEIYEG
301 RCGCKNKDHF LKIDHIGAE KIMEKLPGR EISIQFAGID PIKDPVVPV
351 TTHYMMGGIP TNYHGEVVPV QGDEYEVVK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF RPTPR*

```

g010-1 / P10444

sp|P10444|DHSA_ECOLI SUCCINATE DEHYDROGENASE FLAVOPROTEIN SUBUNIT
 gnl|PID|d1015210 (D90711) Succinate dehydrogenase, flavoprotein [Escherichia coli] gi|1786942
 (AE000175) succinate dehydrogenase flavoprotein subunit [Escherichia coli] Length = 588
 Score = 1073 (495.6 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 191/303 (63%), Positives = 238/303 (78%)

Query: 1 MGFPVRKFDAVIVXXXXXXXXXXXXXSKSLNCVLSKVFPTRSHTVAAQGGISASLGNV 60
 M PVR+FDV++ S+SG CA+LSKVFPTRSHTV+AQGGI+ +LGN
 Sbjct: 1 MKLPVREFDAVIGAGGAGMRAALQISQSGQTCALLSKVFPTRSHTVSAQGGITVALGNT 60

Query: 61 QEDRWDRWHYDTVKGSDDLWQDQDAIEFMCRAPAVIELEHMGMPFDRVESGKIYQRPFG 120
 ED W+WHMYDTVKGSDD++GDQDAIE+MC+ PEA++ELEHMG+PF R++ G+IYQRPFG
 Sbjct: 61 HEDNWEWHYDTVKGSDDYIGDQDAIEYMCKTGPEALLEHMGLPFSRLDDGRIYQRPFG 120

Query: 121 GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRRANTQFFVEWTAQDLIRDENGDDVGV 180
 G + G R A ADRTGHA+LHTLYQQN++ +T F EW A DL++++G VVG
 Sbjct: 121 GQSKNFGGEQAARTAAADRTGHALLHTLYQQNLKNHTTIFSEWYALDLVKNQDGAUVGC 180

Query: 181 TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ 240
 TA+ +ETGEV F A+A + ATGG GRIY S+TNA++NTGDG+G+ RAG+P++DME WQ
 Sbjct: 181 TALCIETGEVYVFKARATVLTGGAGRIYQSTTNAHINTGDGVGMIRAGVPVQDMEMWQ 240

Query: 241 FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDKASRDVVS RAMAMEIYEG 300
 FHPTG+AGAGVL+TEG RCEGG LLN GERFMERYAP KDLA RDVV+R++ +EI EG
 Sbjct: 241 FHPTGIAGAGVLVTEGCRGEGGYLLNKHGERFMERYAPNAKDLAGRDVVARSIMIEIREG 300

Query: 301 RGC 303
 RGC
 Sbjct: 301 RGC 303

Score = 249 (115.0 bits), Expect = 6.7e-169, Sum P(2) = 6.7e-169
 Identities = 53/102 (51%), Positives = 62/102 (60%)

Query: 309 HVLLKIDHIGAEKIMEKLPGIREISIQFAGXXXXXXXXXXXXTTHYMMGGIPTNYHGEV 368
 H LK+DH+G E + +LPGI E+S FA T HYMMGGIPT G+ +
 Sbjct: 310 HAKLKLHLGKEVLESRLPGILELSRTFAHVDVPEKIPVPIPTCHYMMGGIPTKVTGQAL 369

Query: 369 VPQGDEYEVVKGLYAAGECACASVHGANRLGNTNSLLDLVVF 410
 +V V GL+A GE AC SVHGANRLG NSLLDLVVF
 Sbjct: 370 TVNEKGEDVVVPGLFVAGEIACVSVHGANRLGNSLLDLVVF 411

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 69>:

m010-1.seq..

```

1  ATGGGTTTTT CTGTTGCGAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCAGCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTTGTC TAAAGTGTTT CCGACCCGTT CGCATACCGT AGCGGCGCag
151 GGCGGTATTT CCGCCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAAAA TTTATCAGCG
351 TCCTTTCGGC GGCCATACTG CCGAACACGG TAAACGCGCG GTAGAACGCG
401 CCTGTGCGGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC

```

```

451 CAACAAAACG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTTGATT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGGCGGTCG TATTTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCGTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC GGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGGCGCGAG GCGGGTATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTACTGAAAA TCGACCATAT
951 CCGGCGCAGAA AAAATTATGG AAAAAGTCC GGGCATCCGC GAGATTTCCA
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCCG
1051 ACTACCCACT ATATGATGGG CGGCATTCCG ACCAATTACC ACGGCGAAGT
1101 TGTGTTCCG CAAGGTGAAG ATTACGAAGT GCCTGTAATA GGTCTGTATG
1151 CCGCAGGTGA GTGCGCTTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACCAACTCCC TGTGGACTT GGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1301 ATGCAGGTGA GTTACCCGCG CAACGTATCG AGCGTTTGGG CAACCAAACC
1351 GATGGTGAAA ACGTTGATGC ATTGCGTCGC GAACTGCAAC GCTCTGTACA
1401 ACTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GGCATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CCTGATTGAA GTGCGCAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATTCA GATATCAATA CCTGTCTCTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGATGA

```

This corresponds to the amino acid sequence <SEQ ID 70; ORF 010-1>:

m010-1.pep..

```

1 MGFPVRKFDA VIVGGGGAGL RAALQLSKSG LNCVLSKVF PTRSHTVAAQ
51 GGISASLGNV QEDRWDWHMY DTVKGS DWLG DQDAIEFMCRA APEAVIELE
101 HMGMPFDRVE SGKIYQRPFG GHTAEHGKRA VERACAVADR TGHAMLHTLY
151 QQNVRANTQF FVEWTAQDLI RDENG DVVG V TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPT V KDLASRDVVS RAMAMEIYEG
301 RCGCKNKDHF LLKIDHIGAE KIMEKLPGR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVPV QGEDYEVVPK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKP LPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMAIA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPPERDE
551 NWMKHTLYHS DINTLSYKPV HTKPLSVEYI KPAKRVY*

```

m010-1 / g010-1 99.5% identity in 410 aa overlap

	10	20	30	40	50	60
m010-1.pep	MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCVLSKVF PTRSHTVAAQGGISASLGNV					
g010-1	MGFPVRKFDAVIVGGGGAGLRAALQLSKSGLNCVLSKVF PTRSHTVAAQGGISASLGNV					
	10	20	30	40	50	60
m010-1.pep	QEDRWDWHMYDTVKGS DWLG DQDAIEFMCRA APEAVIELEHMGMPFDRVESGKIYQRPFG					
g010-1	QEDRWDWHMYDTVKGS DWLG DQDAIEFMCRA APEAVIELEHMGMPFDRVESGKIYQRPFG					
	70	80	90	100	110	120
m010-1.pep	GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVG V					
g010-1	GHTAEHGKRAVERACAVADRTGHAMLHTLYQQNVRANTQFFVEWTAQDLIRDENG DVVG V					
	130	140	150	160	170	180
m010-1.pep	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
g010-1	TAMEMETGEVYIFHAKAVMFATGGGGRIYASSTNAYMNTGDGLGICARAGIPLEDMEFWQ					
	190	200	210	220	230	240

190

	190	200	210	220	230	240
m010-1.pep	250	260	270	280	290	300
	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVSRAAMEIYEG					
g010-1	250	260	270	280	290	300
	FHPTGVAGAGVLITEGVRGEGGILLNADGERFMERYAPTVDLASRDVVSRAAMEIYEG					
m010-1.pep	310	320	330	340	350	360
	RGCGKNKDHVLLKIDHIGAEMKELPGIREISIQFAGIDPIKDPPIPVVPTTHYMMGGIP					
g010-1	310	320	330	340	350	360
	RGCGKNKDHVLLKIDHIGAEMKELPGIREISIQFAGIDPIKDPPIPVVPTTHYMMGGIP					
m010-1.pep	370	380	390	400	410	420
	TNYHGEVVVPQGEDYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVFEGKAAGDSMIK					
g010-1	370	380	390	400	410	420
	TNYHGEVVVPQGEDYEVPVKGLYAAGECACASVHGANRLGTNSLLDLVVFERTPRX					
m010-1.pep	430	440	450	460	470	480
	FIKEQSDWKPLPANAGELTRQRIERLDNQTDGENVDALRRELQRSVQLHAGVFRTEILS					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 71>:

a010-1.seq..

```

1  ATGGGCTTTC CTGTTGCGAA GTTTGATGCC GTGATTGTCG GCGGTGGTGG
51  TGCAGGTTTA CGCGCANCCC TCCAATTATC CAAATCCGGT CTGAATTGTG
101 CCGTTTGTGC TAAAGTGTTC CCGACCCGTT CGCATACCGT AGCGGCGCAG
151 GGCGGTATT TCCGCTCTCT GGGTAATGTG CAGGAAGACC GTTGGGACTG
201 GCACATGTAC GATACCGTGA AAGGTTCCGA CTGGTTGGGC GACCAAGATG
251 CGATTGAGTT TATGTGCCGC GCCGCGCCTG AAGCCGTAAT TGAGTTGGAA
301 CACATGGGTA TGCCTTTTGA CCGTGTGGAA AGCGGTAATA TTTATCAGCG
351 TCCTTTCGCG GGCCATACTG CCGAACACGG TAAACGCGCG STAGAACGCG
401 CCTGTGCNGT TGCCGACCGT ACAGGTCATG CGATGCTGCA TACTTTGTAC
451 CAACAAAATG TCCGTGCCAA TACGCAATTC TTTGTGGAAT GGACGGCACA
501 AGATTGTGAT CGTGATGAAA ACGGCGATGT CGTCGGCGTA ACCGCCATGG
551 AAATGGAAC CGGCGAAGTT TATATTTTCC ACGCTAAAGC TGTGATGTTT
601 GCTACCGGCG GCGCGGCGCG TATTTATGCG TCTTCTACCA ATGCCTATAT
651 GAATACCGGC GATGGTTTGG GTATTTGTGC GCGTGCAGGT ATCCCCTTGG
701 AAGACATGGA ATTCTGGCAA TTCCACCCGA CCGGCGTGGC AGGTGCGGGC
751 GTGTTGATTA CCGAAGGCGT ACGCGGCGAG GCGGCTATTC TGTGAATGC
801 CGACGGCGAA CGCTTTATGG AACGCTATGC GCCGACCGTA AAAGACTTGG
851 CTTCTCGCGA CGTTGTTTCC CGCGCGATGG CGATGGAAAT CTACGAAGGT
901 CGCGGCTGCG GTAAAAACAA AGACCATGTC TTAAGTAAAA TCGACCATAT
951 CGGCGCAGAA AAAATTATGG AAAAATGACC GGGCATCCGC GAGATTTCCT
1001 TTCAGTTCGC CGGTATCGAT CCGATTAAAG ACCCGATTCC CGTTGTGCGC
1051 ACTACCCACT ATATGATGGG CGGTATTCCG ACCAACTACC ATGGCGAAGT
1101 TGTCGTTTCT CAAGGCGACG AATACGAAGT GCCTGTAAAA GGTCTGTATG
1151 CGGCAGGTGA GTGCGCCTGT GCTTCCGTAC ACGGTGCGAA CCGCTTGGGT
1201 ACGAACTCCC TGCTGGACTT AGTGGTATTC GGTAAAGCTG CCGGCGACAG
1251 CATGATTAAA TTCATCAAAG AGCAAAGCGA CTGGAAACCT TTGCCTGCTA
1301 ATGCCGCGCA ACTGACCCGC CAACGTATCG AGCGTTTGA CAATCAAACCT
1351 GATGGTGAAG ACGTTGATGC ATTGCGCCGC GAACTGCAAC GCTCCGTACA
1401 ATTGCACGCC GCGGTGTTCC GTACTGATGA GATTCTGAGC AAAGGCGTTC
1451 GAGAAGTCAT GCGGATTGCC GAGCGTGTGA AACGTACCGA AATCAAAGAC
1501 AAGAGCAAAG TGTGGAATAC CGCGCGTATC GAGGCTTTGG AATTGGATAA
1551 CTAATTGAA GTGGCGAAAG CGACTTTGGT GTCTGCCGAA GCACGTAAAG
1601 AATCACGCGG TGCGCACGCT TCAGACGACC ATCCTGAGCG CGATGATGAA
1651 AACTGGATGA AACATACGCT GTACCATCA GATGCAATA CCTTGTCTTA
1701 CAAACCGGTG CACACCAAGC CTTTGAGCGT GGAATACATC AAACCGGCCA
1751 AGCGCGTTTA TTGA

```

This corresponds to the amino acid sequence <SEQ ID 72; ORF 010-1.a>:

a010-1.pep..

```

1  MGFFVRKFDA VIVGGGGAGL RAXLQLSKSG LNCVLSKVF PTRSHTVAAQ
51  GGISASLGNV QEDRWDRWHY DTVKGSWLGL DQDAIEFMCR AAPEAVIELE
101  HMGMPPDRVE SGKIIYQRPFG GHATAEHKRA VERACAVADR TGHAMLHTLY

```

151 QQNVRANTQF FVEWTAQDLI RDENGDVVG TAMEMETGEV YIFHAKAVMF
201 ATGGGGRIYA SSTNAYMNTG DGLGICARAG IPLEDMEFWQ FHPTGVAGAG
251 VLITEGVRGE GGILLNADGE RFMERYAPTV KDLASRDVVS RAMAMEIYEG
301 RGCCKNKDHFV LLKIDHIGAE KIMEKLPGIR EISIQFAGID PIKDPIPVVP
351 TTHYMMGGIP TNYHGEVVPV QGDEYEVVPK GLYAAGECAC ASVHGANRLG
401 TNSLLDLVVF GKAAGDSMIK FIKEQSDWKPLPANAGELTR QRIERLDNQT
451 DGENVDALRR ELQRSVQLHA GVFRTEILS KGVREVMIAA ERVKRTEIKD
501 KSKVWNTARI EALELDNLIE VAKATLVSAE ARKESRGAHA SDDHPERDDE
551 NWMKHTLYHS DANTLSYKPV HTKPLSVEYI KPAKRVY*

-- m010-1 / a010-1 99.3% identity in 587 aa overlap

	10	20	30	40	50	60
a010-1.pep	MGFPVRKFD	DAVIVGGGGAGLR	AXLQLSKSGLNCAVLSKVFP	TRSHTVAAQGGISASL	GNV	
m010-1	MGFPVRKFD	DAVIVGGGGAGLR	AXLQLSKSGLNCAVLSKVFP	TRSHTVAAQGGISASL	GNV	
	10	20	30	40	50	60
	70	80	90	100	110	120
a010-1.pep	QEDRW	DWHMYDTVK	GSDWLGDDAIE	FMCRAPPAVIE	LEHMGMPFDR	VESGKIYQRPFG
m010-1	QEDRW	DWHMYDTVK	GSDWLGDDAIE	FMCRAPPAVIE	LEHMGMPFDR	VESGKIYQRPFG
	70	80	90	100	110	120
	130	140	150	160	170	180
a010-1.pep	GHTAEHGKRA	VERACAVADRT	GHAMLHTLYQQ	NVRANTQFFVE	WTAQDLIRD	ENGDVVG
m010-1	GHTAEHGKRA	VERACAVADRT	GHAMLHTLYQQ	NVRANTQFFVE	WTAQDLIRD	ENGDVVG
	130	140	150	160	170	180
	190	200	210	220	230	240
a010-1.pep	TAMEMETGEV	YIFHAKAVMF	FATGGGGRIYAS	SSTNAYMNTG	DGLGICARAGI	PLEDMEFWQ
m010-1	TAMEMETGEV	YIFHAKAVMF	FATGGGGRIYAS	SSTNAYMNTG	DGLGICARAGI	PLEDMEFWQ
	190	200	210	220	230	240
	250	260	270	280	290	300
a010-1.pep	FHPTGVAGAG	VLITEGVRGE	GGILLNADGER	FMERYAPTVK	DLASRDVVS	RAMAMEIYEG
m010-1	FHPTGVAGAG	VLITEGVRGE	GGILLNADGER	FMERYAPTVK	DLASRDVVS	RAMAMEIYEG
	250	260	270	280	290	300
	310	320	330	340	350	360
a010-1.pep	RGCCKNKDHF	VLLKIDHIGAE	KIMEKLPGIRE	ISIQFAGIDP	KDPIPVVPT	THYMMGGIP
m010-1	RGCCKNKDHF	VLLKIDHIGAE	KIMEKLPGIRE	ISIQFAGIDP	KDPIPVVPT	THYMMGGIP
	310	320	330	340	350	360
	370	380	390	400	410	420
a010-1.pep	TNYHGEVVPV	QGDEYEVVPK	GLYAAGECAC	ASVHGANRLG	TNSLLDLVVF	GKAAGDSMIK
m010-1	TNYHGEVVPV	QGDEYEVVPK	GLYAAGECAC	ASVHGANRLG	TNSLLDLVVF	GKAAGDSMIK
	370	380	390	400	410	420
	430	440	450	460	470	480
a010-1.pep	FIKEQSDWKPL	PANAGELTRQ	RIERLDNQT	DGENVDALRR	ELQRSVQLHAG	VFRTEILS
m010-1	FIKEQSDWKPL	PANAGELTRQ	RIERLDNQT	DGENVDALRR	ELQRSVQLHAG	VFRTEILS
	430	440	450	460	470	480
	490	500	510	520	530	540
a010-1.pep	KGVREVMIAA	ERVKRTEIKD	KSKVWNTARIE	ALELDNLIEV	AKATLVSAE	ARKESRGAHA
m010-1	KGVREVMIAA	ERVKRTEIKD	KSKVWNTARIE	ALELDNLIEV	AKATLVSAE	ARKESRGAHA
	490	500	510	520	530	540
	550	560	570	580		
a010-1.pep	SDDHPERDDE	NWMKHTLYHS	DANTLSYKPV	HTKPLSVEYI	KPAKRVYX	
m010-1	SDDHPERDDE	NWMKHTLYHS	DANTLSYKPV	HTKPLSVEYI	KPAKRVYX	

m010-1 SDDHPERDDENWMKHTLYHSDINTLSYKPVHTKPLSVEYIKPAKRVYX
 550 560 570 580

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 73>:

g011.seq
 1 ATGAAGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
 51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
 101 GCCTGAAAAC CCGCCTTACC GAAGATATGA AAACCGCGAT GCGCGCCAAA
 151 GATCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAATGCCG CCGTCAAACA
 201 GTTTGAAGTA GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
 251 TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACGCGCGGAA AATCTACACT
 301 GAAGCCGGCC GTCAGGATTG GGCAGACAAA GAAAACGCCG AAATCGACGT
 351 GCTGCACCGC TACCTGCCGC AAATGCTCTC CGCCGGCGAA ATCCGCACCG
 401 CCGTCGAAGC AGCCGTTGCC GAAACCGGCG CGGCAGGTAT GCGCGATATG
 451 GGCAAAGTGA TGGTCGTATT GAAAaccGC CTCGCCGCA AAGccgATAT
 501 GGGCGAAGTC AACAAAATCT TGAaaaGt aCTGACCGCC tga

This corresponds to the amino acid sequence <SEQ ID 74; ORF 011.ng>:

g011.pep
 1 MKTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKTRLT EDMKTAMRAK
 51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDGAkiYT
 101 EAGRQDLADK ENAEIDLHR YLPQMLSAGE IRTAVEAAVA ETGAAGMADM
 151 GKVMVVLKTR LAGKADMGEV NKILKTVLTA *

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 75>:

m011.seq (partial)
 1 ATGAGGACAC ACCGCAAGAC CTGCTCTGCG GTGTGTTTTG CTTTTCAGAC
 51 GGCATCGAAA CCCGCCGTTT CCATCCGACA TCCCAGCGAG GACATCATGA
 101 GCCTGAAAAT CCGCCTTACC GAAGACATGA AAACCGCGAT GCGCGCCAAA
 151 GACCAAGTTT CCCTCGGCAC CATCCGCCTC ATCAACGCCG CCGTCAAACA
 201 GTTTGAAGTG GACGAACGCA CCGAAGCCGA CGATGCCAAA ATCACCGCCA
 251 TCCTGACCAA AATGGTCAAA CAGCGAAAAG ACAGCGCGAA AATCTACACT
 301 GAAGCCGGCC GTCAGGATTG GGCAGACAAA GAAAACGCCG AAATCGAGGT
 351 ACTGCACCGC TACCTTCCCC AAATGCTTTC CGCCGGCGAA ATCCGTACCG
 401 AGGTCGAAGC TGCCGTTGCC GAAACCGGCG CGGCAGGTAT GCGCGATATG
 451 GGTAAAGTCA TGGGGCTGCT GAAAACCCGC CTCGCAGGTA AAGCCGA...

This corresponds to the amino acid sequence <SEQ ID 76; ORF 011>:

m011.pep (partial)
 1 MRTHRKTCSA VCFAFQTASK PAVSIRHPSE DIMSLKIRLT EDMKTAMRAK
 51 DQVSLGTIRL INAAVKQFEV DERTEADDAK ITAILTKMVK QRKDSAKiYT
 101 EAGRQDLADK ENAEIEVLHR YLPQMLSAGE IRTEVEAAVA ETGAAGMADM
 151 GKVMGLLKTR LAGKA.....

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 011 shows 95.8% identity over a 165 aa overlap with a predicted ORF (ORF 011.ng) from *N. gonorrhoeae*:

m011/g011

	10	20	30	40	50	60
m011.pep	MRTHRKTCSAVCF	AFQTASKPAVSIR	HPSEDIMSLKIR	LTEDMKTAMRAK	DQVSLGTIRL	
g011						
	10	20	30	40	50	60
m011.pep	INAAVKQFEVDER	TEADDAKITAIL	TKMVKQRKDSAK	IYTEAGRQDLAD	KENAEIEVLHR	
g011						
	70	80	90	100	110	120
m011.pep	INAAVKQFEVDER	TEADDAKITAIL	TKMVKQRKDSAK	IYTEAGRQDLAD	KENAEIDLHR	
g011						

193

	70	80	90	100	110	120
	130	140	150	160		
m011.pep	YLPQMLSAGEIRTEVEAAVAETGAAGMADMGKVMGLLKTRLGKA					
g011	YLPQMLSAGEIRTAEEAAVAETGAAGMADMGKVMVVLKTRLGKADMGEVNNKILKTVLTA					
	130	140	150	160	170	180
g011	X					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 77>:

g012.seq

```

1  ATGCTCGCCC GTCGCTATTT TTTCAATATC CAACCCGGGG CGGTTTTCAC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGCCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACa
201 gGcggTGGAT ATTCGgcact tccgCcacca caccaccga accgatgacc
251 gcaaacggaG CGGAAACAAT TTTATCCGcC acacacgcca tcatatagcc
301 gcCGCTTGCC GCGACCTTAT CGAcggcgac ggTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CTAAGCTGC GAAGCCGCCA AACCCTAACC GTGAACCACG
401 CCGCCCGGAC TTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAAATCGCC GTCAAAATCC AACACAAAAA GCGGGGATTT TTGCGTTTCG
551 GCAGATTCTC CCCCGCCCTC CTTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CTTTTTTTTC CTGATGTTTT GTCTCTTCCT
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 78; ORF 012.ng>:

g012.pep

```

1  MLARRYFFNI QPGAVFTDKL LEQLMRFLQF LPEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRHFRHHTHR TDDRKRSGNN FIRHTRHHIA
101 AACRDLDIGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS EQNLIFRLGN
151 QKHRRNLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPAL LQTLFLCFGF
201 RLFLFLFFFF LMFCLFLA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 79>:

m012.seq

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCCGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GGCGGTGGAT ATTCGGTACT TCCGCCACCA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 GCCGCTCGCn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
351 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
401 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
451 nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn nnnnnnnnnnnn
501 nnnnnnnnnnnn nnnnnnnnnnnC AACACAAAAA GCGGTGATTT nTGC GTTTTCG
551 GCAGATTCTC CCCACCCCTC CTTCAAACGT TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTGT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 80; ORF 012>:

m012.pep

```

1  MLARCHFLNI QIRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRYFRHHTHR TDNRKRSGSN FIRHTRHHIT
101 AARXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
151 XXXXXXXXXXX XXXXXXXXXXX XXXQHKKA*F XRFGRFLPTL LQTFFLCFGF
201 RLFLFLFLFF LMLCLFPA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 81>:

```
a012.seq
1   ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GCGGTGGGAT ATTCGGTACT TCCGCTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CCTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACACG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTT TTGCGTTTCG
551 GAAGATTTCT CCCACCCCTC CTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA
```

This corresponds to the amino acid sequence <SEQ ID 82; ORF 012.a>:

```
a012.pep
1   MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQA VD IRYFRYNTHR TDNRKRSGNN FIRHTRHHIT
101 TARRHLIDGD GQRNIAFAQT PKLRSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKRRRLMTQ GFYGVCIQIA VKIQHKKAGF LRFGRFLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFCCLFPA*
```

m012/a012 64.2% identity over a 218 aa overlap

	10	20	30	40	50	60
m012.pep	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
a012	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
	10	20	30	40	50	60
m012.pep	70	80	90	100	110	120
	NIMFFQQAVIDIRYFRHHTHRTDNRKRSGSNFIRHTRHHITAAARXXXXXXXXXXXXXXXXXX					
a012	NIMFFQQAVIDIRYFRYNTHRTDNRKRSGNNFIRHTRHHITTARRHLIDGDGQRNIAFAQT					
	70	80	90	100	110	120
m012.pep	130	140	150	160	170	180
	XXXQHKKAXF					
a012	PKLRSRQTVTVNHAARTFQSKQNLIFRLGNQKRRRLMTQGFYGVCIQIAVKIQHKKAGF					
	130	140	150	160	170	180
m012.pep	190	200	210	219		
	XRFGRLPTLLQTLFLCFGRFLFLFLFFLMCLFPAX					
a012	LRFGRLPTLLQTLFLCFGRFLFLFLFFLMFCCLFPAX					
	190	200	210			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 012 shows 58.7% identity over a 218 aa overlap with a predicted ORF (ORF 012.ng) from *N. gonorrhoeae*:

m012/g012

	10	20	30	40	50	60
m012.pep	MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
g012	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					

195

	10	20	30	40	50	60
m012.pep	70	80	90	100	110	120
	NIMFFQQAVDIRYFRHHTHRTDNRKSGSNFIRHTRHHITAARXXXXXXXXXXXXXXXXXXXX					
g012	: : : : : :					
	70	80	90	100	110	120
	NIMFFQQAVDIRHFRHHTHRTDDRKSGNNFIRHTRHHIAAACRDLDGDGQRNIAFAQT					
	130	140	150	160	170	180
m012.pep	XXXQHKKAXF					
	:	:	:	:	:	:
g012	PKLRSRQTVTVNHAARTFQSEQNLIIFRLGNQKHRRLMTQGFYGVCIQIAVKIQHKKAGF					
	130	140	150	160	170	180
	XRFGRFLPTLLQTFFLCFGRFLFLFLFLFLMLCLFPAX					
m012.pep	: : : : : :					
g012	LRFGFRFLPALQLTFLCFGRFLFLFLFLFLMFLCLFLAX					
	190	200	210	219		
	XRFGRFLPTLLQTFFLCFGRFLFLFLFLFLMLCLFPAX					
	: : : : : :					
	LRFGFRFLPALQLTFLCFGRFLFLFLFLFLMFLCLFLAX					
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 83>:

m012-1.seq

```

1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGGTTCTCGC
51  TGACAAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCGGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GCGGTGGAT ATTCCGTACT TCCGCCACCA CACCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAGCAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 GCCGCTCGCC GCCACCTTAT CGACGCGGAC GGTCAGCGGA ATATTGCGTT
351 CGCGCAAACG CyTAAGCTGC GAAGCCGCCA AACCCTAACG GTGAACCAG
401 CCGCCCGGAC TTTCCAATCT GAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAAATCGCC GTCAAATCC AACACAAAAA GGCGGGATT TCGGTTTCG
551 GCAGATTCTC CCCCACCCTC CTTCAAACGC TTTTCTCTG CTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 84; ORF 012-1>:

m012-1.pep

```

1  MLARCHFLNI QLRVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRYFRHHTH RDNKRSNFI RHRHHTH
101 AARRHLIDGD GQRNIAFAQT XKLRSRQTVT VNHAARTFQ SEQNLIIFRLGN
151 QKHRRLMTQ GFYGVCIQIA VKIQHKKAGF LRFGFRFLPTL LQTLFLCFGF
201 RLFLFLFLFF LMFLCLFPA*

```

m012-1/g012 91.7% identity in 218 aa overlap

	10	20	30	40	50	60
m012-1.pep	MLARCHFLNIQLRAVLADKLLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI					
g012	MLARRYFFNIQPGAVFTDKLLEQLMRFLQFLPEFLFALFRIFTHKSNRALKFARRHHIHI					
	10	20	30	40	50	60
	NIMFFQQAVDIRYFRHHTHRTDNRKSGSNFIRHTRHHITAARRHLIDGDGQRNIAFAQT					
m012-1.pep	: : : : : :					
g012	NIMFFQQAVDIRHFRHHTHRTDDRKSGNNFIRHTRHHIAAACRDLDGDGQRNIAFAQT					
	70	80	90	100	110	120
	XKLRSRQTVTVNHAARTFQSEQNLIIFRLGNQKHRRLMTQGFYGVCIQIAVKIQHKKAGF					
m012-1.pep	: : : : : :					
g012	PKLRSRQTVTVNHAARTFQSEQNLIIFRLGNQKHRRLMTQGFYGVCIQIAVKIQHKKAGF					
	130	140	150	160	170	180

196

	130	140	150	160	170	180
	190	200	210	219		
m012-1.pep	LRFG	RFLP	TLQL	TLFL	CFGR	FLFLFLFFLMFCLFPAX
		:				
g012	LRFG	RFLP	ALLQ	TLFL	CFGR	FLFLFLFFLMFCLFLAX
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 85>:

```

a012-1.seq
1  ATGCTCGCCC GTTGCCACTT CCTCAATATC CAATTGAGGG CGTTTCTCGC
51  TGACAACTG CTTGAACAAC TGATGCGTTT CCTCCAGTTC CTGTCCGAAT
101 TTCTGTTTGC CCTTTTCCGT ATTTTCACCC ATAAAAGTAA CCGTGCGCTT
151 AAATTCGCCC GCCGTCATCA CATCCACATC AATATCATGT TTTTCAACA
201 GGCGGTGGAT ATTCGGTACT TCCGTACAA CACCCACCGA ACCGACAATC
251 GCAAACGGAG CGGAAACAAT TTTATCCGCC ACACACGCCA TCATATAACC
301 ACCGCTCGCC GCCACCTTAT CGACGGCGAC GGTGAGCGGA ATATTGCGTT
351 CGCGCAAACG CTAAGCTGC GAAGCCGCCA AACCGTAACC GTGAACACAG
401 CCGCCCGGAC TTTCCAATCT AAGCAGAACC TCATCTTCAG GCTTGGCAAT
451 CAAAAGCACC GCCGTAATCT CATGACGCAA GGATTCTACG GCGTGTGCAT
501 ACAATCGCC GTCAAAATCC AACACAAAAA GGCGGGATTG TTGCGTTTCG
551 GAAGATTTCT CCCACCCCTC CTCAAACGC TTTTCTCTG CTTTGGCTTC
601 CGCCTTTTCC TTTTCTTTT CCTCTTTTC CTGATGTTT GCCTCTTCCC
651 CGCTTAA

```

This corresponds to the amino acid sequence <SEQ ID 86; ORF 012-1.a>:

```

a012-1.pep
1  MLARCHFLNI QLRAVLADKL LEQLMRFLQF LSEFLFALFR IFTHKSNRAL
51  KFARRHHIHI NIMFFQQAVD IRYFRYNTHR TDNRKSGNN FIRHTRHHIT
101 TARRHLIDGD GQRNIAFAQT PKLSRQTVT VNHAARTFQS KQNLIFRLGN
151 QKHRRNLTQ GFYGVCIQIA VKIQHKKAGF LRFGRLPTL LQTLFLCFGF
201 RLFLFLFLFF LMECLFPA*

a012-1/m012-1 97.2% identity in 218 aa overlap

a012-1.pep
10 20 30 40 50 60
MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI
|||||
m012-1
10 20 30 40 50 60
MLARCHFLNIQLRAVLADKLEQLMRFLQFLSEFLFALFRIFTHKSNRALKFARRHHIHI

a012-1.pep
70 80 90 100 110 120
NIMFFQQAVDIRYFRYNTHRTDNRKSGNNFIRHTRHHITTARRHLIDGDGQRNIAFAQT
|||||
m012-1
70 80 90 100 110 120
NIMFFQQAVDIRYFRHHTHRTDNRKSGNSFIRHTRHHITAARRHLIDGDGQRNIAFAQT

a012-1.pep
130 140 150 160 170 180
PKLSRQTVT VNHAARTFQSKQNLIFRLGNQKHRRNLTQGFYGVCIQIAVKIQHKKAGF
|||||
m012-1
130 140 150 160 170 180
XKLSRQTVT VNHAARTFQSEQNLIFRLGNQKHRRNLTQGFYGVCIQIAVKIQHKKAGF

a012-1.pep
190 200 210 219
LRFGRLPTLLQTLFLCFGRFLFLFLFLFFLMFCLFPAX
|||||
m012-1
190 200 210
LRFGRLPTLLQTLFLCFGRFLFLFLFLFFLMFCLFPAX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 87>:

```

g013.seq
1  aTgcctttga ccatgctgtg cagcaGGAcg tGCGGTTtgt tcataatata
51  gtCcgaccGG AAAagcggAG GAAaCGCAGT GCCGCGCCCT TCCCCTTTCT
101 TGCCGTGGCA GCGATGCag tTgGATTCTG ACACCTTTTG CCCTTTtGtc

```

197

```

151 atgatGCTgt tgtcggCGGC AGAAGCgGCG GcgCAGAGGC AGCACAAGAT
201 GAAGGCGGTC GGCAGTCGGG TTGTGTtcat tGgcgTTTCC cctaattgttt
251 tgaaaccttg ttttttgatt Ttgcctttac ggggtgaaaa gtttttTtgg
301 cccaaatccg gaatttag

```

This corresponds to the amino acid sequence <SEQ ID 88; ORF 013.ng>

g013.pep

```

1  MPLTMLCSRT CGLFIQSDR KSGGNAVPRP SPFLPWQAMQ LDSYTFCPFV
51  MMLLSAAEAA AQRQHKMKAV GSRVVFIVGS PNVLKPCFLI LPLRGEKFFW
101 PKSGI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 89>:

m013.seq

```

1  ATGCCTTTGA CCATGCTGTG CAGCAGCACC TGCGGTTTTT TCATGATGAA
51  GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
101 TGCCGTGGCA GGCAGCGCAG TTGGATTTCG AACTTTTTTG CCCTTTTGTC
151 ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAAGC AGCCCAAGAC
201 GAGGGCGGTC GGCAGTCGGG TTGTGTTTCAT TGGTGTTC TCATGTTTG
251 AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT TTGCGCGAAT
301 CAGTAG

```

This corresponds to the amino acid sequence <SEQ ID 90; ORF 013>:

m013.pep

```

1  MPLTMLCSST CGFFMMKSER XSGGNMVRP SPFLPWQATQ LDSYTFCPFV
51  MMLLSAAEAA AQKQPKTRAV GSRVVFIVGS FMFETLLLIL RSGXKIFLPN
101 Q*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 91>:

a013.seq

```

1  ATGCCTTTGA CCATGCTGTG CAGCAGCACC TGCGGTTTTT TCATGATGAA
51  GTCGGAGCGG TAGAGCGGCG GAAACATGGT TCCGCGGCCT TCGCCCTTTT
101 TGCCGTGGCA GGCAGCGCAG TTGGATTTCG AACTTTTTTG CCCTTTTGTC
151 ATGATGCTGT TGTCGGCGGC AGAAGCGGCG GCGCAGAGGC AGCCCAAGAC
201 GAGGGCGGTC GGCAGTCGGG TTGTGTTTCAT TGGTGTTC TTAATGTTTG
251 AAACCTTGTT GTTGATTTTG CGTAGCGGGT GAAAGATTTT CTTGCCGAAT
301 CCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 92; ORF 013.a>:

a013.pep

```

1  MPLTMLCSST CGFFMMKSER *SGGNMVRP SPFLPWQATQ LDSYTFCPFV
51  MMLLSAAEAA AQKQPKTRAV GSRVVFIVGS LMFETLLLIL RSG*KIFLPN
101 R*

```

m013/a013 97.0% identity over a 101 aa overlap

	10	20	30	40	50	60
m013.pep	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTFCPFVMMMLLSAAEAA					
a013	MPLTMLCSSTCGFFMMKSERXSGGNMVRPSPFLPWQATQLDSYTFCPFVMMMLLSAAEAA					
	10	20	30	40	50	60
	70	80	90	100		
m013.pep	AQKQPKTRAVGSRVVFIVGSFMFETLLILRSGXKIFLPNQX					
	:					
a013	AQKQPKTRAVGSRVVFIVGSLMFETLLILRSGXKIFLPNRX					
	70	80	90	100		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 013 shows 73.3% identity over a 101 aa overlap with a predicted ORF (ORF 013.ng) from *N. gonorrhoeae*:

	10	20	30	40	50	60
m013.pep	MPLTMLCSSTCGFFMMKSERXSGGNMVRPRSPFLPWQATQLDSYTCFPMMLLSAAEAA					
	: : : :					
g013	MPLTMLCSRTCGLFI IQSDRKSGGNAVPRSPFLPWQAAMQLDSYTCFPVVMMLLSAAEAA					
	10	20	30	40	50	60
	70	80	90	100		
m013.pep	AQKQPKTRAVGSRVVFIGVSF-MFETLLILR-SGXKIFLPNQX					
	: : : : : : :					
g013	AQRQHMKAVGSRVVFIGVSPNVLKPCFLLPLRGKKFFWPKGSGIX					
	70	80	90	100		

g015.seq

1	ATGCAGTATC	TGATTGTCAA	ATACAGCCAT	CAATCTTCG	TTACCATCAC
51	CATTTTGGTA	TTCAACATCC	GTTTTTCCT	ACTTTGGAAA	AATCCAGAAA
101	AGCCCTTTGGT	CGGCTTTTGG	AAGCACTACG	CCCACCTCAA	CGACACGATG
151	CTGCTGTTTA	CGGGATTGTG	GCTGATGAAG	ATTACCAATT	TCTCCCCTTT
201	CAACGCGCCT	TGGCTCGGCA	CAAAAATCCT	GCTCCTGTTC	GCCTACATCG
251	CACCTGGGCAT	GGTAATGATG	CGCGCCCGTC	CGCGTTCGAC	CAAGTTCTAC
301	ACCGTTTACC	TGCTCGCTAT	TGTGTCGATC	GCCTGCATCG	TTTACCTTGC
351	CAAAACCAAA	GTCCTGCCAT	TCGTA		

g015.pap

1 MQYLIVKYSH QIFVTITILV FNIRFLLWK NPEKPLVGFW KALPHLNDTM
51 LLFTGLWLWK ITHFSPFNAP WLGTKILLF AYIALGMVMM RARPRSTKFY
101 TVYLLAMCCI ACIVYLAKTK VLPF*

m015.seq

```

1      . . AAAATCAGAA AAGCCTTGGC GGGCTTTTGG AAGGCACTGC CCCACCTTAA
51     CGACACCATG CTGCTGTTTA CGGGATTGTG GCTGATGAAA ATTACCCATT
101    TCTCCCGTGT CAACGCGCCT TGGCTCGGTA CAAAATCCT GCTTCTGTC
151    GCCTATATCT CATTTGGGTAT GATGATGATG CGCGCCCGTC CGCGTTTCGAC
201    CAAGTTCAC  ACCGTTTACC TGCTGCCCAT GTGTTGCGTC GCCTGCATCG
251    TTTACCTTGC CAAAACCAAA GTCCTGCCTT TCTGA

```

m015.pep (partial)

1 ..KIRKALAGFW KALPHLNDTM LLETGLWLMK ITHFSPFNAP WLGTKILLLL
51 AYIALGMMM RARPRSTKFY TVYLLAMCCV ACIVYLAKTK VLPF*

a015.seq

1	ATGCAGTATC	TGATTGTCAA	ATACAGCCAT	CAATCTTCG	TTACCATCAC
51	CATTTTGGTA	TTCAACATCC	GTGTTTNCNT	ACTTTGGAAA	AATCCAGAAA
101	AGCCCCTTGGC	GGGCTTTTGG	AAGGCACTGA	CCCCACTTAA	CGACACCATG
151	CTGCTGTTTA	CGGGATTGTG	GCTGATGAAA	ATTACCACATT	TCTCCCGGTT
201	CAACGCGCCT	TGGCTCGGTA	CAAAAATCCT	GCTTCTGCTC	GCCTATATCG
251	CATTGGGTAT	GATGATGATG	CGCGCCCGTC	CGGTTGAC	CAAGTTCTAC
301	ACCGTTTACC	TGCTGCAGAT	GTGTTGCCCT	ACCTGCATCG	TTACCTTGC
351	CAAAACAAA	GTCTGCGCTT	TCTGA		

a015.pap

199

1 MQYLIVKYSH QIFVTITILV FNIRVFXLWK NPEKPLAGFW KALPHLNDTM
 51 LLFTGLWLMK ITHFSPFNAP WLGTKILLLL AYIALGMMM RARPRSTKFY
 101 TVYLLAMCCL TCIVYLAKTK VLPF*

m015/a015 96.7% identity over a 91 aa overlap

m015.pep				10	20	30
				KIRKALAGFWKALPHLNDT	MLLFTGLWLMKITH	
a015	LIVKYSHQIFVTITILVFNIRVFXLWKNPEKPLAGFWKALPHLNDT	MLLFTGLWLMKITH				
	10	20	30	40	50	60
m015.pep	40	50	60	70	80	90
	FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTK	FYTVYLLAMCCVACIVYLAKTKVLP				
a015	FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTK	FYTVYLLAMCCLTCIVYLAKTKVLP				
	70	80	90	100	110	120
m015.pep	FX					
a015	FX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 015 shows 94.5% identity over a 91 aa overlap with a predicted ORF (ORF 015.ng) from *N. gonorrhoeae*:

m015/g015

m015.pep				10	20	30
				KIRKALAGFWKALPHLNDT	MLLFTGLWLMKITH	
g015	LIVKYSHQIFVTITILVFNIRFLLWKNPEKPLVGF	WALPHLNDT	MLLFTGLWLMKITH			
	10	20	30	40	50	60
m015.pep	40	50	60	70	80	90
	FSPFNAPWLGTKILLLLAYIALGMMMRRARPRSTK	FYTVYLLAMCCVACIVYLAKTKVLP				
g015	FSPFNAPWLGTKILLLLFAYIALGMVMMRRARPRSTK	FYTVYLLAMCCVACIVYLAKTKVLP				
	70	80	90	100	110	120
m015.pep	FX					
g015	FX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 99>:

g018.seq

1 atGCAGCAGG GGCagttggt tggacgcgtc gcccgcaata AAGATATGCG
 51 GAATgctggt CTGCATggtC AGCGGATCGG CAACGGGtac gccgcgcgcg
 101 tctttgTCGA TATTGATGTT TTCCAAACCG ATATgtTCAA CGTTCCGGACG
 151 GCgACCTACG GCTGCCAACA TATATTCGGC AACAAATACG CCTTTTTCGC
 201 CATCTGCTC CCAATGGACT tctACATTGC CGTCTGCGTC GAGTTTGACC
 251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT tctTCTCCGA ACACGGCTTT
 301 CGCCTCGTCT GA

This corresponds to the amino acid sequence <SEQ ID 100; ORF 018.ng>:

g018.pep

1 MQQGQLVGRV ARNKDMRNAG LHGQRIGNGY AARVFVDIDV FQTDIVNVRT
 51 ATYGCQHIFG NKYAFAFILL PMDFYIAVCV EFDLGFSIQM QFQFFSEHGF

200

101 RLV*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 101>:

```
m018.seq
1  ATGCAGCAGA GGCAGTTGGT TGGACGCATC GCCTGCGATG AAGATATGCG
51  GAATACTGGT CTGCATGGTC AGCGGGTCGG CAACAGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATATT TTCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCATC GAGTTTGACC
251 TCGGTTTTAG CATCCAGATG CAGTTTCAAT TCTTCGCCGA ACACGGCGTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 102; ORF 018>:

```
m018.pep
1  MQQRQLVGRI ACDEDMRNTG LHGQRVGNRY AARIFFDIDI FQTDIVNVRT
51  AAHGCGHIFG NKYAFAILL PMDFYIAVCI EFDLGFSIQM QFQFFAEHGV
101 RLV*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 103>:

```
a018.seq
1  ATGCAGCAGG GGCAGTTGGT TGGACGCGTC GCCCGCAATA AAGATATGCG
51  GAATACTGGT CTGCATAGTC AGCGGATCGG CAACGGGTAC GCCGCGCGCA
101 TCTTTTTCGA TATTGATGTT TTCAAACCG ATATTGTCAA CGTTCGGACG
151 GCGGCCACG GCTGCCAGCA TATATTCGGC AACAAATACG CCTTTTTCGC
201 CATCCTGCTC CCAATGGACT TCTACATTGC CGTCTGCGTC GAGTTTGACC
251 TCGGTTTTAG CATCAAATG CAGTTTCAAT TCTTCACCGA ACACGGCTTT
301 CGCCTCGTCT GA
```

This corresponds to the amino acid sequence <SEQ ID 104; ORF 018.a>:

```
a018.pep
1  MQQGQLVGRV ARNKDMRNTG LHSQRIGNGY AARIFFDIDV FQTDIVNVRT
51  AAYGCGHIFG NKYAFAILL PMDFYIAVCV EFGLGFSIQM QFQFFTEHGF
101 RLV*
```

m018/a018 86.4% identity over a 103 aa overlap

	10	20	30	40	50	60
m018.pep	MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCGHIFG					
			:	:	:	:
a018	MQQGQLVGRVARNKDMRNTGLHSQRIGNGYAARIFFDIDVFQTDIVNVRTAAYGCGHIFG					
	10	20	30	40	50	60
	70	80	90	100		
m018.pep	NKYAFAILLPMDFYIAVCI EFDLGFSIQM QFQFFAEHGVRLVX					
			:	:	:	:
a018	NKYAFAILLPMDFYIAVCV EFGLGFSIQM QFQFFTEHGFRLVX					
	70	80	90	100		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 018 shows 84.5% identity over a 103 aa overlap with a predicted ORF (ORF 018.ng) from *N. gonorrhoeae*:

m018/g018

	10	20	30	40	50	60
m018.pep	MQQRQLVGRIACDEDMRNTGLHGQRVGNRYAARIFFDIDIFQTDIVNVRTAAHGCGHIFG					
			:	:	:	:
g018	MQQGQLVGRVARNKDMRNAGLHGQRIGNGYAARVFVDIDVFQTDIVNVRTATYGCQHIFG					
	10	20	30	40	50	60

201

	70	80	90	100
m018.pep	NKYAFFAILLPMDFYIAVCIEFDLGFSIQMVFQFFAEHGVRLVX			
	: : :			
g018	NKYAFFAILLPMDFYIAVCVEFDLGFSIQMVFQFFSEHGFRVLX			
	70	80	90	100

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 105>:

```

g019.seq (partial)
1  ..ctgctggcgg ccctgggtget tgccgcgtgt tcttcgACAA ACacacTGCC
51  AGCCGGCAAG ACCCCGGCAG ACAATATAGA AActgcCgAC CTTTCGGCAA
101 GCGTTCCAC cgcCCTGCC GAACCGAAG GAAAAACGCT GGCAGATTAC
151 GGCGGCTACC CGTCCGCACT GGATGCAGTG AAACAGAACA ACGATGCGGC
201 AGCCGCCGCC TATTTGGAAÀ Acgcaggaga cagCGcgatg gcGAAAatg
251 tccgcaagga gtgGCTGa

```

This corresponds to the amino acid sequence <SEQ ID 106; ORF 019.ng>:

```

g019.pep (partial)
1  ..LLAALVLAAC SSTNTLPAGK TPADNIETAD LSASVPTRPA EPEGKTLADY
51  GGYPSALDAV KQNNDAAAAA YLENAGDSAM AENVRKEWL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 107>:

```

m019.seq (partial)
1  ATGTACCTAC CCTCTATGAA GCATTCCCTG CCGCTGCTGG CGGCCCTGGT
51  GCTTGCCGCG TGTCTTTCGA CAAACACACT GCCAGCCGGC AAGACCCCGG
101 CAGACAATAT AGAACTGCC GACCTTTCGG CAAGCGTTCC CACCCGCCCT
151 GCCGAACCCG AAAGAAAAAC GCTGGCAGAT TACGGCGGGT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCCGCAA CGAGTGGCTG
301 AAGTCTTTGG GCGCACGCG ACAGTGGACG CTGTTTGAC ACGAATACGC
351 CAAACTCGAA CCGGCAGGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCGCTG AACTGGTCAA AAATACGGGC
451 AAACGTGCCTT CGGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGC GCGGA CTGCTGGCCG
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
601 TTTGACGGCG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAAGAA GCACGCAAAT CGCCGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGTG GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCTT TGGACTATTA
801 CGGCAAGGTT GCCGACCGCC GCCAATGAC CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCCTTGCGC GCCCGACGTT GGGACGAGCT GGCCTCCGTT
901 ATCTCGCATA TGCCGAAAA ACTGCAAAAA AGCCCGACCT GGCTCTACTG
951 GCTGGCACGC AGCCGCGCGC CAACGGGCAA CACGCAAGAG GCGGAAAAAC
1001 TTTACAAACA GGCGGCAGCG ACGGGCAGGA ATTTTATGC GGTGCTGGCA
1051 GGGGAAGAAT TGGGTCGGA AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAAAC AGCGTCCGCC GCATGGCGGA AGACGGTGCA GTCAAACGCG
1151 CACTGGTACT GTTCCAAAAC AGCCAATCTG CCGGTGATGC AAAAATGCGC
1201 CGTCAGGCTC AGGCGGAATG GCGTTTGGC ACACGCGGCT TTGACGAAGA
1251 CAAGCTGCTG ACCGCCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCGACC GCAAACCTCA CTACACCTTG
1351 CGCTATATTT CGCCGTTTAA AGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCTTGGG TTTATGGGCT GATTGCTCAG GAAAGCCGCT
1451 TCGTTATAGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1501 ATGCCTGCCA CCGCGCGCGA AATCGCCGCG AAAATCGGTA TGGATGCCGC
1551 ACAACTTTAC ACCGCCGACG GG...

```

This corresponds to the amino acid sequence <SEQ ID 108; ORF 019>:

```

m019.pep (partial)
1  MYLPSMKHSL PLLAALVLA CSSTNTLPAG KTPADNIETA DLSASVPTRP

```

```

51  AEPERKTLAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARQWT LFAQEYAKLE PAGRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWVGL
251 GHYQSQNLNV PAALDYXGKV ADRRLTDDQ IEWYARAALR ARRWDELASV
301 ISHMPEKLQK SPTWLYWLAR SRAATGNTQE AEKLYKQAAA TGRNIFYAVLA
351 GEELGRKIDT RNNVPDAGKN SVRRMAEDGA VKRALVLFQN SQSAGDAKMR
401 RQAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNLYTL
451 RYISPFKDTV IRHAQNVNVD PAWVYGLIRQ ESRFVIGAQS RVGAQGLMQV
501 MPATAREIAG KIGMDAAQLY TADG...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 109>:

```

a019.seq
1  ATGTACCCAC CCTCTCTGAA GCATTCCCTG CCGCTGCTGG TGGNCCTGGT
51  GCTTGCCGCG TGTCTTNGA CAAACACACT GTCAGCCGAC AAGACCCCGG
101 CAGACAATAT AGAAACTGCC GACCTTTCGG CAAGCGTCC CACCNGCCCT
151 GCCGAACCCG AANGAAAAAC GTNGGCAGAT TACGGCGGCT ACCCGTCCGC
201 ACTGGATGCA GTGAAACAGA AAAACGATGC CGCCGTCGCC GCCTATTGG
251 AAAACGCCGG CGACAGCGCG ATGGCGGAAA ATGTCGCGAA CGAGTGGCTG
301 AAGTCTTTGG CCGCGCGCAG ACAGTGGACG CTGNTGCAC ANGAATATGC
351 NAAACTCGAA CCGGCANGGC GCGCCCAAGA AGTCGAATGC TACGCCGATT
401 CGAGCCGCAA CGACTATACG CGTGCCCGCG AACTGGTCAA AAATACGGGC
451 AAAGTGCCTT CCGGCTGCAC CAAACTGTTG GAACAGGCAG CCGCATCCGG
501 CTTGTTGGAC GGCAACGACG CCTGGAGGCG CGTGCGCGGA CTGCTGGCGC
551 GCCGCCAAAC CACAGACGCA CGCAACCTTG CCGCCGCATT GGGCAGCCCG
601 TTTGACGCGG GTACACAAGG TTCGCGCGAA TATGCCCTGT TGAACGTCAT
651 CGGCAAGAA GCACGCAAT GCCTGAATGC CGCCGCCCTG CTGTCCGAAA
701 TGGAAAGCGG TTTAAGCCTC GAACAACGCA GTTTCGCGT GGGCGTATTG
751 GGGCATTATC AGTCGCAAAA CCTCAATGTG CCTGCCGCT TGGACTATTA
801 NGGCAAGGTT GCCGACCGCC GCCAAGTACG CGACGACCAA ATCGAGTGGT
851 ACGCCCGCGC CGCNNNTNGC NNNCGNNGT NGNANGANN TGGCNCNGNN
901 ANCNCGNNNN TGCNNGANAA ACNNNNNNAN AGCNCNANN NGNTNNANTG
951 NNTGGCACGC AGCCGCGCCG CNACGGGCAA CACGCAANAN GCGGANAAAC
1001 TNTACAAACA GGCGGCAGCA NCGGGCANGA ATTTTATATG NGTGCTGNCN
1051 GGGGAAGAGT TGGGGCGCAN AATCGATACG CGCAACAATG TGCCCGATGC
1101 CGGCAAAANC AGCGTCTCC GTATGGCGGA AGACGGCGCG ATTAAGCGCG
1151 CGCTGGTGCT GTTCCGAAAC AGCCGAACCG CCGCGCATGC GAAAATGCGC
1201 CGTCNGGCTC AGCGGGAATG GCGTTTCGCC ACACGCGGCT TCGATGAAGA
1251 CAAGCTGCTG ACCGCCGCGC AAACCGCGTT CGACCACGGT TTTTACGATA
1301 TGGCGGTCAA CAGCGCGGAA CGCACCAGAC GCAAACCTCA CTACACCTTG
1351 CGCTACATTT CGNNNNNTNA NGACACGGTA ATCCGCCACG CGCAAAATGT
1401 TAATGTCGAT CCGGCGTGGG TTTACGGGCT GATTCGTCAG GAAAGCCGCT
1451 TCGTTATGGG CGCGCAATCC CGCGTAGGCG CGCAGGGGCT GATGCAGGTT
1501 ATGCCTGCCA CCGCGCGCGA AATCGCCGGC AAAATCGGTA TGGATGCCGC
1551 ACAACTTTAC ACCGCCGACG GCAATATCCG TATGGGGACG TGGTATATGG
1601 CGGACACCAA ACGCCGCTG CAAAACAACG AAGTCCTCGC CACCGCAGGC
1651 TATAACGCCG GTCCCGGCAG GGCGCGCCGA TGGCAGGCGG ACACGCCCTT
1701 CGAAGGCGCG GTATATGCCG AAACCATCCC GTTTTCCGAA ACGCGCGACT
1751 ATGTCAAAAA AGTGATGGCC AATGCCGCCT ACTACGCCTC CCTCTTCGGC
1801 GCGCCGCACA TCCCGCTCAA ACAGCGTATG GGCATTGTCC CCGCCGCTG
1851 A

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 019.a>:

```

a019.pep
1  MYPPSLKHSL PLLVXLVLA CSXTNTLSAD KTPADNIETA DLSASVPTXP
51  AEPFXKTXAD YGGYPSALDA VKQKNDAAVA AYLENAGDSA MAENVRNEWL
101 KSLGARQWT LXAXEYAKLE PAXRAQEVEC YADSSRNDYT RAAELVKNTG
151 KLPSGCTKLL EQAAASGLLD GNDAWRRVRG LLAGRQTDDA RNLAALGSP
201 FDGGTQGSRE YALLNVIGKE ARKSPNAAAL LSEMESGLSL EQRSFAWVGL
251 GHYQSQNLNV PAALDYXGKV ADRRLTDDQ IEWYARAAXX XRXXXXXXAXX
301 XXXXXXXXXX XXXXXXXXAR SRAATGNTQX AXKLYKQAAA XGXNIFYAVLX
351 GEELGRXIDT RNNVPDAGKV SVLRMAEDGA IKRALVLFNR SRTAGDAKMR
401 RXAQAEWRFA TRGFDEDKLL TAAQTAFDHG FYDMAVNSAE RTDRKLNLYTL

```


451 RYISXXXTV IRHAQNVNVD PAWVYGLIRQ ESRFVMAQS RVGAQGLMQV
 501 MPATAREIAG KIGMDAAQLY TADGNIRMGY WYADTKRRL QNNEVLATAG
 551 YNAGPGRARR WQADTLEGA VYAETIPFSE TRDYVKKVMA NAAYYASLFG
 601 APHILPKQRM GIVPAR*

m019/a019 88.9% identity over a 524 aa overlap

m019.pep	10	20	30	40	50	60
	MYLPSMKHSLPLLAALVLAACSSTNTLPAGKTPADNIETADLSASVPTRPAPERKTLAD					
a019	10	20	30	40	50	60
	MYPPSLKHSLPLLVLVLAACSXTNTLSADKTPADNIETADLSASVPTXPAPERKTXAD					
m019.pep	70	80	90	100	110	120
	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKS LGARRQWTLFAQEYAKLE					
a019	70	80	90	100	110	120
	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKS LGARRQWTLXAXEYAKLE					
m019.pep	130	140	150	160	170	180
	PAGRAQEVECYADSSRNDYTRAAELVKNTGKLP SGCTKLLEQAAASGLLDGNDAWRRVRG					
a019	130	140	150	160	170	180
	PAXRAQEVECYADSSRNDYTRAAELVKNTGKLP SGCTKLLEQAAASGLLDGNDAWRRVRG					
m019.pep	190	200	210	220	230	240
	LLAGRQTTDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEMESGLSL					
a019	190	200	210	220	230	240
	LLAGRQTTDARNLAAALGSPFDGGTQGSREYALLNVIGKEARKSPNAAALLSEMESGLSL					
m019.pep	250	260	270	280	290	300
	EQRSAWGVLGHYQSQNLNVPALDYYGKVADRRQLTDDQIEWYARAALRARRWDELASV					
a019	250	260	270	280	290	300
	EQRSAWGVLGHYQSQNLNVPALDYYGKVADRRQLTDDQIEWYARAAXXXXXXXXAXX					
m019.pep	310	320	330	340	350	360
	ISHMPEKLQKSPTWLYWLARSRAATGNTQEAELKYQAAATGRNFYAVLAGEELGRKIDT					
a019	310	320	330	340	350	360
	XXXXXXXXXXXXXXXXXARSRAATGNTQXAKLYQAAAXGXNFYAVLXGEELGRXIDT					
m019.pep	370	380	390	400	410	420
	RNNVPDAGKNSVRRMAEDGAVKRALVLFQNSQSAGDAKMRRQAQAEWRFATRGFDEDKLL					
a019	370	380	390	400	410	420
	RNNVPDAGKXSVLRMAEDGAIKRALVLFNRNRTAGDAKMRRQAQAEWRFATRGFDEDKLL					
m019.pep	430	440	450	460	470	480
	TAAQTAFDHGFYDMAVNSAERTDRKLNYYTLRYISPFKDTVIRHAQNVNVDPAWVYGLIRQ					
a019	430	440	450	460	470	480
	TAAQTAFDHGFYDMAVNSAERTDRKLNYYTLRYISXXXTVIRHAQNVNVDPAWVYGLIRQ					
m019.pep	490	500	510	520		
	ESRFVIGAQS RVGAQGLMQVMPATAREIAGKIGMDAAQLYTADG					
a019	490	500	510	520	530	540
	ESRFVIGAQS RVGAQGLMQVMPATAREIAGKIGMDAAQLYTADGNIRMGY WYADTKRRL					
a019	550	560	570	580	590	600
	QNNEVLATAGYNAGPGRARRWQADTLEGA VYAETIPFSE TRDYVKKVMA NAAYYASLFG					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 019 shows 95.5% identity over a 89 aa overlap with a predicted ORF (ORF 019.ng) from *N. gonorrhoeae*:

g019/m019

		10	20	30	40	49
g019.pep		LLAALVLAACSSNTLPAGKTPADNIETADLSASVPTRPAEPEGKTLAD				
m019		MYLPSMKHSLPLAALVLAACSSNTLPAGKTPADNIETADLSASVPTRPAPERKTLAD				
		10	20	30	40	50
	50	60	70	80	89	
g019.pep	YGGYPSALDAVKQNNDAAYLENAGDSAMAENVRKEWL					
m019	YGGYPSALDAVKQKNDAAVAAYLENAGDSAMAENVRNEWLKS LGARROWTLFAQEYAKLE					
	70	80	90	100	110	120

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 111>:

g023.seq

```

1   ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GCGATTGGGT
51  AATGCAGCGT GCGACTGCGG TTATTATGTT GATTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTGCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATT
151 TTTAGTCAAG CTTGGGTAAA AGTATTTACC CAAGTGAGCT TTATCGCCGT
201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCCTTCGG CGTGCGTTTG TTTTGCAGG TTGCCACCAT TGtctGGCTG
301 GTCGGCTGCC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 023.ng>:

g023.pep

```

1   MVERKLGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFAI PKEYPAWQAF
51  FSQAWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

m023.seq

```

1   ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GCGATTGGGT
51  GATGCAACGT GCGACTGCGG TTATTATGTT GATTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTCCCTG CCTAAAGAAT ATCCGGCATG GCAGGCATT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT
201 ATTCTTGCAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATCA
251 AACCCCTTCGG CGTGCGTTTG TTTTGCAGG TTGCCACCAT CGTTTGGCTG
301 GTCGGCTGTC TCGTGTATTC AGTTAAAGTG ATTTGGGGGT AA

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 023>:

m023.pep

```

1   MVERKLGAH YGLRDWVMQR ATAVIMLIYT VALLVVLFSI PKEYSAWQAF
51  FSQFWVKVFT QVSFIAVFLH AWVGIRDLWM DYIKPFGVRL FLQVATIVWL
101 VGCLVYSVKV IWG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

a023.seq

```

1   ATGGTAGAAC GTAAATTGAC CGGTGCCCCAT TACGGTTTGC GGGATTGGGC
51  GATGCAACGT GCGACCGCGG TTATTATGTT GATTATACC GTTGCACTTT
101 TAGTGGTTCT ATTTGCTCTG CCTAAAGAAT ATCCGGCATG GCAGGCATT
151 TTTAGTCAAA CTTGGGTAAA AGTATTTACC CAAGTGAGCT TCATCGCCGT

```

201 ATTCTTGAC GCTTGGGTGG GTATCCGCGA TTTGTGGATG GACTATATNA
 251 AACCCCTTCGG CGTGCGTTTG TTTTGCAGG TTGCCACCAT CGTCTGGCTG
 301 GTCGGCTGCT TGGTGTATTC AATTAAAGTA ATTTGGGGGT AA

This corresponds to the amino acid sequence <SEQ ID 116; ORF 023.a>:

a023.pep
 1 MVERKLTGAH YGLRDWAMQR ATAVIMLIYT VALLVVLFA L PKEYSAWQAF
 51 FSQTWVKVFT QVSFIAVFLH AWVGIRDLWM DYXKPFVRL FLQVATIVWL
 101 VGCLVYSIKV IWG*

m023/a023 96.5% identity over a 113 aa overlap

	10	20	30	40	50	60
m023.pep	MVERKLTGAHYGLRDWVMQRATAVIMLIYT VALLVVLFSLPKEYSAWQAFFSQTWVKVFT					
	: : : : :					
a023	MVERKLTGAHYGLRDWAMQRATAVIMLIYT VALLVVLFA L PKEYSAWQAFFSQTWVKVFT					
	10	20	30	40	50	60
	70	80	90	100	110	
m023.pep	QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSVKVIWGX					
	: : : :					
a023	QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSIKVIWGX					
	70	80	90	100	110	

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 023 shows 97.3% identity over a 113 aa overlap with a predicted ORF (ORF 023.ng)

from *N. gonorrhoeae*:

g023/m023

	10	20	30	40	50	60
g023.pep	MVERKLTGAHYGLRDWVMQRATAVIMLIYT VALLVVLFA L PKEYPAWQAFFSQAWVKVFT					
	: : : :					
m023	MVERKLTGAHYGLRDWVMQRATAVIMLIYT VALLVVLFSLPKEYSAWQAFFSQTWVKVFT					
	10	20	30	40	50	60
	70	80	90	100	110	
g023.pep	QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSVKVIWGX					
	: : : :					
m023	QVSFIAVFLHAWVGIRDLWMDYIKPFGVRLFLQVATIVWL VGCLVYSVKVIWGX					
	70	80	90	100	110	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 117>:

g025.seq

1 ATGTTGAAAC AAacgACACT TTTGGCAGCT TGTACCGCCG TTGCCGCTCT
 51 GTTGGGCGGT TGcgCCACCC AACAGCCTGC TccTGTCATT GCAGGCAATT
 101 CAGGTATGCA GACCGTATCG TCTGCGCCGG TTTACAATCC TTATGGCGCA
 151 ACGCCGTACA ATGCCGCTCC TGCCGCCAac gatgcGCCgT ATGTGCCGCC
 201 CGTGCAAact gcgcgggttT ATTCGCCTCC TGCTTATGTT CCGCcgTCTG
 251 CACCTGCCGT TTCGGgtaca tatgtTCCTT CTTACGCACC CgtcgACATC
 301 aacgCGGCGa cgCataCTAT TGTGCGTGGC GACACgGtgt acaACATTTC
 351 caaAcgCtac CATATCTCTC AAGACGATTT CCGTGCCTGG AACGGCATGA
 401 CCGACAATAC GTTGAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCaggA
 451 TATGCCGCAC CGAAAACCGC AGCCGTAGAA AGCAGGCCCG CCGTACCGGC
 501 TGCCGCGCAA ACCCTGTGA AACCCGCCGC gcaACCGCCC GTTCAGTCCG
 551 CGCCGCAACC TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCCGCCCCC
 601 GCGCCCGCCC CGCAATCTCC TGCCGCTTCG CCTTCCGGCA CGCGTTCGGT
 651 CGGCGGCATT GTTTGGCAGC GTCCGACCCA AGGTAAAGTG GTTGCCGATT

```

701 TCGGCGGCGG CAACAAGGGT GTCGATATTG CCGCAATGC CGGACAACCC
751 GTTTTGGCGG CGGCTGACGG CAAAGTGGTT TATGCCGGTT CAGGTTTGAG
801 GGGATACGGA AACTTGGTCA TCATCCAGCA CAATTCCTCT TTCCTGACCG
851 CGTACGGGCA CAACCAAAAA TTGCTGGTCG GCGAAGGTCA GCAGGTCAAA
901 CGCGGTCAGC AGGTGCTTTT GATGGGTAAT ACCGATGCTT CCAGAACGCA
951 GCTTCATTTT GAGGTGCGTC AAAACGGCAA ACCGGTTAAC CCGAACAGCT
1001 ATATCGCGTT CTGA

```

... This corresponds to the amino acid sequence <SEQ ID 118; ORF 025.ng>:

g025.pep

```

1 MLKQTLLAA CTAVAALLGG CATQQPAPVI AGNSGMQTVS SAPVYNPYGA
51 TPNNAAPAAAN DAPYVPPVQT APVYSPPAYV PPSAPAVSGT YVPSYAPVDI
101 NAATHTIVRG DTVYNISKRY HISQDDFRAW NGMTDNTLSI GQIVKVKPAG
151 YAAPKTAAVE SRPAVPAAQ TPVKPAAQPP VQSAPQPAAP AAENKAVPAP
201 APAPQSPAAS PSGTRSVGGI VWQRPTQKQV VADFGGNGK VDIAGNAGQP
251 VLAADGKVV YAGSGLRGY NLVIIQHNSS FLTAYGHNQK LLVGEGQQVK
301 RGQVALMGN TDASRTQLHF EVRQNGKPVN PNSYIAF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

m025.seq (partial)

```

1 ..GTGCCGCGG TGCAAAGCGC GCCGGTTTAT ACGCCTCCTG CTTATGTTCC
51 GCCGTCTGCA CCTGCCGTTT CGGGTACATA CGTTCCTTCT TACGCACCCG
101 TCGACATCAA CGCGCGGACG CATACTATTG TGCGCGGCGA CACGGTGTAC
151 AACATTTCCA ACGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGA
201 CGGCATGACC GACAATACGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC
251 CGGCAGGATA TGCCGCACCG AAAGCCGCAG CCGTAAAAAG CAGGCCCGCC
301 GTACCGGCTG CCGCGCAACC GCCCGTACAG TCCGCACCCG TCGACATTAA
351 CGCGGCGACG CATACTATTG TGCGCGGCGA CACGGTGTAC AACATTTCCA
401 AAGCTACCA TATCTCTCAA GACGATTTCC GTGCGTGGA CCGCATGACC
451 GACAATATGT TGAGCATCGG TCAGATTGTT AAAGTCAAAC CGGCAGGATA
501 TGCCGCACCG AAAACCGCAG CCGTAGAAAG CAGGCCCGCC GTACCGGCTG
551 CCGTGCAAAC CCCTGTGAAA CCCGCCGCGC AACCGCCTGT GCAGTCCGCG
601 CCGCAACCTG CCGCGCCCGC TGCGGAAAAT AAAGCGGTTT CCGCGCCCGC
651 CCGCAATCT CCTGCCGCTT CGCCTTCCGG CACGCGTTTC GTGCGCGGCA
701 TTGTTTGGCA GCGTCCGACG CAAGGTAAAG TGGTTGCCGA TTTCGGCGGC
751 AACAACAAGG GTGTCGATAT TGCCGGTAAT GCGGGACAGC CCGTTTGGC
801 GCGGCTGAC GGCAAAGTGG TTTATGCCGG TTCAGGTTT AGGGGATACG
851 GAACTTGGT CATCATCCAG CATAATTTCT CTTTCCTGAC CGCATACGGG
901 CACAACCAA AATTGCTGGT CGGCGAGGGG CAGCAGGTCA AACGCGGTCA
951 GCAGGTGCT TTGATGGGCA ATACCGATGC TTCCAGAACG CAGTTTCATT
1001 TCGAGGTGCG TCAAACGGC AAACCGGTTA ACCCGAACAG CTATATCGCG
1051 TTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 025>:

m025.pep (partial)

```

1 ..VPPVQSAPVY TPPAYVPPSA PAVSGTYVPS YAPVDINAAT HTIVRGDTVY
51 NISKRYHISQ DDFRAWNGMT DNTLSIGQIV KVKPAGYAAP KAAAVKSRPA
101 VPAAQPPVQ SAPVDINAAT HTIVRGDTVY NISKRYHISQ DDFRAWNGMT
151 DNLSIGQIV KVKPAGYAAP KTAAVESRPA VPAAVQTPVK PAAQPPVQSA
201 PQPAAPAAEN KAVPAPAPQS PAASPSGTRS VGGIVWQRPT QGKVVADFGG
251 NNKGVDIAGN AGQPVLAAD GKVVYAGSGL RGYGNLVIIQ HNSSFLLTAYG
301 HNKKLLVGEG QQVKGQQVA LMGNTDASRT QLHFEVRQNG KPVNPNSYIA
351 F*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

a025.seq

```

1 ATGTTGACAC CAACAACACT TTAGGTAGCT TGTACCGCCC TTGCCGCTCA
51 GTTGGGCGGA TGCCCCACCC AACACCCTTC TCCTGTCATT GCAGGCAATT
101 CAGGTATGCA GACCGTACCG TCTGCGCCGG TTTACAATCC TTATGGCGCA

```

207

```

151  ACGCCGTACA ATGCCGCTCC TGCCGCCAAC GATGCGCCGT ATGTGCCGCC
201  GGTGCAAAGC GCGCCGGTTT ATANGCCTCC TGCTTATGTT CCGCCGTCTG
251  CACCTGCCGT TTCGGGTACA TACGTTCTT CTTACGCANC CGTCGACATC
301  AACCGCGCGA CCCATACTAT TGTGCGCGGC GACACCGTGT ACAAGATTTT
351  CAAATGCTAC CATATCTCTC AAGACGATT CCGTGCGTGG AACGGCATGA
401  CCGACAATAC GTTGAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCAGGA
451  TATGCCGCAC CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC
501  TGCCGCGCAA CCGCTCGTAC AGTCCGCACC CGTCGACATC AACCGCGCGA
551  CGCATACTAT TGTGCGCGGC GACACGGTGT ACAACATTTC CAAACGCTAC
601  CATATCTCTC AAGACGATT CCGTGCGTGG AACGGCATGA CCGACAATAC
651  GTTGAGCATC GGTGAGATTG TTAAAGTCAA ACCGGCAGGA TATGCCGCAC
701  CGAAAGCCGC AGCCGTAAAA AGCAGGCCCG CCGTACCGGC TGCCGTGCAA
751  ACCCCTGTGA AACCGCGCGC GCAACCGCCT GTGCAGTCCG CGCCGCAACC
801  TGCCGCGCCC GCTGCGGAAA ATAAAGCGGT TCCGCGCCCC GCCCGCAAT
851  CTCCTGCCCG TTCGCTTCC GGCACGCGTT CGGTCGGCGG CATTTGTTGG
901  CACGCTCCGA CGCAAGGTAA AGTGTTGCC GATTTCGGCG GCAACAACAA
951  GGGTGTCGAT ATTGCAGGAA ATGCGGGACA GCGCGTTTGT GCGGCGGCTG
1001 ACGGCAAAGT GGTATTATGCA GGTTCGGGT TGAGGGGATA CGGCAATTTG
1051 GTCATCATCC AGCATAATTC TTCCTTCTG ACCGCATACG GGCACAACCA
1101 AAAATTGCTG GTCGCGGAAG GCCAGCAGGT CAAACGCGGG CAGCAGGTCG
1151 CTTTGATGGG CAATACCGAG GCTTCTAGAA CGCAGCTTCA TTTCGAGGTG
1201 CGGCAAAACG GCAAACCGGT TAATCCGAAC AGCTATATCG CGTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 025.a>:

```

a025.pep
  1  MLTP TTL*VA CTALAAQLGG CPTQHPSPIV AGNSGMQTVS SAPVYNPYGA
 51  TPYNAA PAAN DAPYVFPVQS APVYXPPAYV PPSAPAVSGT YVPSYAXVDI
101  NAATHI VRG DTVYKISKCY HISQDD FRAW NGMTDNTLSI GQIVKVKPAG
151  YAAPKAA AVK SRPAVFAAAQ PLVQSAPVDI NAATHI VRG DTVYNISKRY
201  HISQDD FRAW NGMTDNTLSI GQIVKVKPAG YAAPKAA AVK SRPAVFAAVQ
251  TPVKPAA QPP VQSAPQPAAP AAENKAVPAP APQSPAASPS GTRSVGGIVW
301  QRPTQGKVVA DFGGNNKGVD IAGNAGQPV LAAADGKVVA GSGLRGYGNL
351  VIIQHNS SFL TAYGHNQKLL VEGEQQVKRG QQVALMGNT EASRTQLHFEV
401  RQNGKPVNPN SYIAF*

m025/a025 97.4% identity over a 351 aa overlap

m025.pep
                                     10      20      30
                                     VPPVQSAPVYTPPAYVPPSAPAVSGTYVPS
                                     |||||:|||||
a025      GMQTVPSAPVYNPYGATPYNAAPAANDAPYVPPVQSAPVYXPPAYVPPSAPAVSGTYVPS
              40      50      60      70      80      90

m025.pep
              40      50      60      70      80      90
YAPVDINAATHI VRGDTVYNISKRYHISQDD FRAWNGMTDNTLSIGQIVKVKPAGYAAP
|| |||||:|||||
a025      YAXVDINAATHI VRGDTVYKISKCYHISQDD FRAWNGMTDNTLSIGQIVKVKPAGYAAP
              100     110     120     130     140     150

m025.pep
              100     110     120     130     140     150
KAAAVKSRPAVFAAAQPPVQSAPVDINAATHI VRGDTVYNISKRYHISQDD FRAWNGMT
|||||
a025      KAAAVKSRPAVFAAAQPLVQSAPVDINAATHI VRGDTVYNISKRYHISQDD FRAWNGMT
              160     170     180     190     200     210

m025.pep
              160     170     180     190     200     210
DNMLSIGQIVKVKPAGYAAPKTAAVESRPAPVPAVQT PVKPAAQPPVQSAPQPAAPAAEN
|| |||||:|||||
a025      DNTLSIGQIVKVKPAGYAAPKAAVKS RPAPVPAVQT PVKPAAQPPVQSAPQPAAPAAEN
              220     230     240     250     260     270

m025.pep
              220     230     240     250     260     270
KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGVVADFGGNNKGVDIAGNAGQPVLAAD
|||||

```

208

```

a025      KAVPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLAAD
           280      290      300      310      320      330

m025.pep      280      290      300      310      320      330
GKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEQQVVRGQQVALMGNTDASRT
|||||:|||||:|||||:|||||:|||||:|||||
a025      GKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEQQVVRGQQVALMGNTDASRT
           340      350      360      370      380      390

m025.pep      340      350
QLHFEVRQNGKPVNPNSYIAFX
|||||
a025      QLHFEVRQNGKPVNPNSYIAFX
           400      410

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 025 shows 75.6% identity over a 353 aa overlap with a predicted ORF (ORF 025.ng) from *N. gonorrhoeae*:

m025/g025

```

m025.pep      10      20      30
VPPVQSAPVYTTPPAYVPPSAPAVSGTYVPS
|||||:|||||:|||||:|||||:|||||
g025      GMQTVSSAPVYNPYGATPYNAAPAANDAPYVPPVQTAPVYSPPAYVPPSAPAVSGTYVPS
           40      50      60      70      80      90

m025.pep      40      50      60      70      80      90
YAPVDINAATHITVRGDTVYNISKRYHISQDDFRANGMTDNTLSIGQIVKVPAGYAAP
|||||:|||||:|||||:|||||:|||||
g025      YAPVDINAATHITVRGDTVYNISKRYHISQDDFRANGMTDNTLSIGQIVKVPAGYAAP
           100     110     120     130     140     150

m025.pep      100     110     120     130     140     150
KAAAVKSRPAVPAAQPPVQSAPVDINAATHITVRGDTVYNISKRYHISQDDFRANGMT
|
g025      K-----

m025.pep      160     170     180     190     200     210
DNMLSIGQIVKVPAGYAAPKTAAVESRPAPVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
|||||:|||||:|||||:|||||:|||||
g025      -----TAAVESRPAPVPAVQTPVKPAAQPPVQSAPQPAAPAAEN
           160     170     180     190

m025.pep      220     230     240     250     260
KAVPAPAP--QSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLA
|||||:|||||:|||||:|||||:|||||
g025      KAVPAPAPAPQSPAASPSGTRSVGGIVWQRPTQGKVVADFGGNNKGVDIAGNAGQPVLA
           200     210     220     230     240     250

m025.pep      270     280     290     300     310     320
ADGKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEQQVVRGQQVALMGNTDAS
|||||:|||||:|||||:|||||:|||||
g025      ADGKVVYAGSGLRGYGNLVIIQHNSSFLTAYGHNQKLLVGEQQVVRGQQVALMGNTDAS
           260     270     280     290     300     310

m025.pep      330     340     350
RTQLHFEVRQNGKPVNPNSYIAFX

```

g025 |||||
 RTQLHFEVRQNGKPVNPNSYIAFX
 320 330

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 113>:

g031.seq
 1 ATGGTGTCCC TCCGCTTCAG ATTCCGCAAC CACTTTAAAC GCCGACATTC
 51 TGACAATTTT CTTTTCGCGC AGCCAAATAT CATGCGTATC TTTCGGTTCG
 101 GGCTTGTGG GCATGGCAAC CTTCAACAGC CGCGCCATEA CAGGAATCGT
 151 CGTTCCCTGA ATCAGCAGCG ACAGCACCAC CACGGCAAAC GCCACATCAA
 201 ACAGCAGGTG CGAATTGGGA ACGCCCATCA CCAGCGGCAT CATCGCCAGC
 251 GAAATCGGTA CGGCTCCTCG CAAGCCCAAC CAACTGATAT ACGCCTTTTC
 301 ACGCAGGCTG TAATTGAATT TCCACAAACC GCCGAACACT GCCAGCGGAC
 351 GCGCGACCAG CATCAGGAAC GCCGCAATCG CCAAGGCTTC CGCCGCCCTG
 401 TCCAACACGC CGGCGGGAGA AACCAGCAGA CCGAGCATGA CGAACAAAGT
 451 TGCTGCGCGC AGCCAAGCCA AACCGTCCAT CACACGCAA ACGTGTTCG
 501 TcgACGGTT GCGCTGGTTA CCGACAATGA TGCCGGCAAG GTAAACCGCC
 551 AAAAAGCCGC TGC GCCTAT GGTATTGTA AACGCAAACA CAAGCAGCCC
 601 GCCCGACACA ATCATCAGCG CGTACAGACC TTCCGtacac acctccaatt
 651 cccaatcaac gtcatactgt tctccctgt taaaatgttc ttcacttcag
 701 aatccccccc ttcttccag cccgaaacct tcatgtgtta naccctgggg
 751 tgccccaacg gatttagtaa cctcccaatg actctgcttg tcgccccctt
 801 cgcccgcttt ctcttccg gaaaacttgt tgtccccgtc ttacattaa

This corresponds to the amino acid sequence <SEQ ID 114; ORF 031.ng>:

g031.pep
 1 MVSRLRFRGN HFKRRHSDNF LFRQPNIMRI FRFGLVGHGN LQQPRHHRNR
 51 RSLNQQRQHH HGKRHIKQV RIGNAHHQRH HRQRNRYGSS QAQPTDIRLF
 101 TQAVIEFPQT AEHCQRTDQ HQERRNRQGF RRPVQHAGGR NQQTEHDEQS
 151 CLRQPSQTVH HTQNVFRRTV ALVTDNDAGK VNRQKAAAY GIGKRKHQKQ
 201 ARHNHQRVQT FRTHLQFPIN VIAVSRVKMF FTSESPSSQ PETFMCXTLG
 251 CPNGFSNLPM TLLVAPFARF LLPGKLVVPV LH*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 115>:

m031.seq (partial)
 1 ...CGCCTGAAGC ACGGTGTCGG ACTGCATTTC TATTCGGCTA TACGCCTTTT
 51 CACGCAGGCT GTAATTGAAT TTCCACAAAC CGCCGAACAC TGCCGACGGA
 101 CGCGCGACCA GCATCAGGAA CGCCGCAATC GCCAAGGCTT CGCCCGCCTT
 151 GTCCAACACG TTGGCAGGAG AAACCAGCAG CAAAGGCATT CCCAACGTG
 201 CGGACAAAGT GGTGAAACC ACGCTCAGAA ACAACAGTGC GCCACCCGGC
 251 AG....

This corresponds to the amino acid sequence <SEQ ID 116; ORF 031>:

m031.pep (partial)
 1 ...RLKHGVLHF YSAIRLFTQA VIEFPQTAEH CRRTRDQHQE RRRNRQGFRRP
 51 VQHVGRNRQQ QRHSQTCGQS GRNHAQKQOC ATRQ....

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

a031.seq
 1 ATACGCCTTT TCACGCAGGC TGTAATTGAA TTTCCACAAA CCGCCGAACA
 51 CTGCCGCGCG ACGCGCGACC AGCATCAGGA ACGCCGCAAT CGCCAAGGCT
 101 TCCGCCGCCC CGTCCAACAC GTTGGCAGGA GAAACCAGCA GCAAAGGCAT
 151 TCCCAAACGT GCGGACAAAG TGGTCGAAAC CACGCTCAGA AACAACAGTG
 201 CGCCACCCGG CAG

This corresponds to the amino acid sequence <SEQ ID 118; ORF 031.a>:

a031.pep (partial)
 1 IRLFTQAVIE FPQTAEHCRR TRDQHQERRN RQGFRRPVQH VGRNRQQQRH
 51 SQTGQSGRN HAQKQCATR Q

210

m031/a031 100.0% identity over a 71 aa overlap

```

          10      20      30      40      50      60
m031.pep  RLKHGVLHGFYSAIRLFTQAVIEFPQTAEHCRRTRDQHQERRNRQGFRRPVQHVGRNRQQ
          |||||
a031      IRLFTQAVIEFPQTAEHCRRTRDQHQERRNRQGFRRPVQHVGRNRQQ
          10      20      30      40

          70      80
m031.pep  QRHSQTCGQSGRNHAQKQCCATRQ
          |||||
a031      QRHSQTCGQSGRNHAQKQCCATRQ
          50      60      70

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 031 shows 60.0% identity over a 85 aa overlap with a predicted ORF (ORF 031.ng) from *N. gonorrhoeae*:

m031/g031

```

          10      20      30
m031.pep  RLKHGVLHGFYSAIRLFTQAVIEFPQTAEH
          |::|: : |||||
g031      NQQRQHHHGKRHIKQQVRIGNAHQHRRQRNRYGSSQAQPTDIRLFTQAVIEFPQTAEH
          60      70      80      90      100     110

          40      50      60      70      80
m031.pep  CRRTRDQHQERRNRQGFRRPVQHVGRNRQQQRHS-QTCGQSGRNHAQKQCCATRQ
          |:|||||:|||||:| ||| :|: |:: ::: | : |:
g031      CQRTDQHQERRNRQGFRRPVQHVGRNRQQTEHDEQSCLRQPSQTVHHTQNVFRTVALV
          120     130     140     150     160     170

g031      TDNDAGKVNRRQKAAAYGIGKRKHKQPARHNRQVQTFRTHLQFPINVIASRVKMFFTS
          180     190     200     210     220     230

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 119>:

g032.seq

```

1  ATGCGGCGAA ACGTGCCTGC CGTCGCCGTA TTGCGCCGCC CACGATTCTGA
51  GGCCTTTTGT GATTGGCGGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGCTTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAAGGCTTCC ACGCTTTTGC CGGTACGCGG AACCTGACGC TGCTTGCGCC
201 CTTTGCCGGT AACGTGTACC CACGCTTCGT CCAAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAAA CCGCTGCCGT ACATCAGCTC
301 GAACAGCGCG TGGTCGCGCA CCGCCAGCGG GTCGCGCCG TCCACGGGCA
351 AATCCAACAT CCGGTTACAG CATTCTCTGC GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGCGGTTT GATGTCGGCG GTCGGGTCGG CGCGCATCAG
451 CCCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCACGCC GACAGCTTGC
501 GCGCCAGCGT CCGACCGTCC AAACCGCGCT GCGACAGCCG CCGCAACGCC
551 GccgTAAAT CGCGCCGCGA CAAGTCCTGC GGCACGCcgc ctgcaTCTTC
601 AGACGGCATT TGTGCCAACA GTGCAAACAG TTCTTCCAAA TCGCGCCGGT
651 ATGCCGCAAC CGTGTGCTCC GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701 TAAGCGTCAA AATacgcccG AAACccgTCC AAAACCATAA CCGTCCACAA
751 CAAATATCAA AAAACAGTG A

```

This corresponds to the amino acid sequence <SEQ ID 110; ORF 032.ng>:

g032.pep

```

1  MRRNVPAVAV LRRPRFEAFI DLALAQARAV PAGKQGFVAV CRTLQRQIVF
51  QGFHAFAGQR NLTLLAPFAG NVYPREFVOIY IICIQAVYLA HAQTAAVHQL

```


211

```

101 EQRVVAHRQR VAAVHGQIQH PVQPFRLQGF GYALGLLRRF DVGGRVGAHQ
151 PAFDQPGAIL PPRRLARQR PTVQTLARQP PQRRLKIAPR QVLRHAACIF
201 RRHLCQQCKQ FFQIAPVCRN RVLRLALAHQ VFQISVKIRR KPVQNHNRPT
251 QISKNQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 111>:

```

m032.seq (partial)
1 ATGCGGCGAA ACGTGCmTGC mGTCGCCGTT kTGCGCCGCC CATTGCGCCA
51 AACGTTTTTG GATTGCGGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGACCAGCGG CACCTGCCGC TgTT.GCGCC
201 CTTTGCCGAT AAcGTGTACC CACGCyTCGT CCAATAGAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAA CCGCTGCCGT ACATCAGTTC
301 GAACAGGGCG TGGTCGCGCA CCGCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTGAGC CATTCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGGCGTTT GATGTCGGCG GTCGGGTCGG CGTGATCAG
451 GCCGCGCTTT ACCAGCCAAA CGCAATACTG CCGCCAAGAC GAAAGCTTGC
501 GAGCCAGCGT CCGTTCCCCC AAACCGCG...

```

This corresponds to the amino acid sequence <SEQ ID 112; ORF 032>:

```

m032.pep (partial)
1 MRRNVXAVAV XRRPLRQTFL DLALAQARAV PAGKQGFVR CRLTQRQIVF
51 QGFHAFADQR HLPLXAFAD NVYPRXVQID IICIQAVYLA HAQTAAVHQF
101 EQGVVAHRQR VAAVHGQIQH PVQPFRLQGF GYALGLLRRF DVGGRVGVHQ
151 AALYQPNAIL PPRRLASQR FFPQTA...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 113>:

```

a032.seq
1 ATGCGGCGAA ACGTGCCTGC CGTCGCCGTT TTGCGCCGCC CATTGCGCCA
51 AACGTTTTTG GATTGCGGT TGGCTCAGGC GCGTGCCGTT CCTGCCGGTA
101 AACAGGGCTT TGCCGTCCGA TGCCGTCTGA CGCAGCGGCA GATAGTTTTT
151 CAGGGCTTCC ACGCTTTTGC CGGTCAGCGG AACCTGCCGC TGCTTGCCTC
201 CTTTGCCGGT AACGTGTACC CACGCCTCGT CCAATATAC ATCATCTGCA
251 TTCAAGCCGT GTATCTCGCT CACGCGCAA CCGCTGCCGT ACATCAGTTC
301 GAACAGCGCG TGATCGCGCA CCGCAGCGG GTCGCCGCCG TCCACGGGCA
351 AATCCAGCAT CCGGTTGAGC CATTCTGCG GCAGGGCTTT GGGTACGCGC
401 TCGGGCTGCT TCGGGCGTTT GATGTCGGCG GTCGGGTCGG TATGCAGCAG
451 ACCGCGTTTG ACCAGCCAGG CGCAATACTG CCGCCAAGAC GACAGCTTGC
501 GCGCCAGCGT CCGCGCATT CAAACCGCGT GCGACAGCCG CCGCAACGCC
551 CCGGTAAAAT CGCGCTGCGA CAAGCCCTGC GGCACGCCGC CTGCATCTTC
601 AGACGGCATT TGTGCCAACA GCGCAAACAG TTCTTCCAAA TCGCGCCGGT
651 ATGCCGCCAC CGTGCTGCT GACTTGCCCT CGCGCACGAT GTTTTCCAAA
701 TAAGCGTCAA AATGCGCCGC AAACCCGTCC AAAACCATAA CCGCCCCACA
751 CAAATATCAA AAAACAGTG A

```

This corresponds to the amino acid sequence <SEQ ID 114; ORF 032.a>:

```

a032.pep
1 MRRNVPAVAV LRRPLRQTFL DLALAQARAV PAGKQGFVR CRLTQRQIVF
51 QGFHAFAGQR NLPLASFAG NVYPRLVQIY IICIQAVYLA HAQTAAVHQF
101 EQRVIAHRQR VAAVHGQIQH PVQPFRLQGF GYALGLLRRF DVGGRVGMQQ
151 TAFDQPGAIL PPRRLARQR PRIQTALRQP PQRRLKIALR QALRHAACIF
201 RRHLCQQRKQ FFQIAPVCRH RVLRLALAHQ VFQISVKMRR KPVQNHNRPT
251 QISKNQ*

```

m032/a032 88.1% identity over a 176 aa overlap

	10	20	30	40	50	60
m032.pep	MRRNVXAVAVXRRPLRQTFLDLALAQARAVPAGKQGFVRCRLTQRQIVFQGFHAFADQR					
a032	MRRNVPAVAVLRRPLRQTFLDLALAQARAVPAGKQGFVRCRLTQRQIVFQGFHAFAGQR					
	10	20	30	40	50	60

m032/g032

	10	20	30	40	50	60
m032.pep	MRRNVXAVAVXRRPLRQTFLDLALAQAARAVPAGKQGFVAVRCRLTQRQIVFQGFHAFADQR					
	:					
g032	MRRNVPAVAVLRRPRFEAFDLDLALAQAARAVPAGKQGFVAVRCRLTQRQIVFQGFHAFAGOR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m032.pep	HLPLXAPFADNVYPRXVQIDIIICIQAVYLAHAQTAAVHQFEQGVVAHRQRVAAVHQGIQH					
	:			:		
g032	NLTLLAPFAGNVYPRFVQIYIICIQAVYLAHAQTAAVHQLEQRVVAHRQRVAAVHQGIQH					
	70	80	90	100	110	120
	130	140	150	160	170	
m032.pep	PVQPFLLRQGFYALGLLRRFDVGGRVGVHQAALYQPNAILPERRKLASQRPFQTA					
				:	:	
g032	PVQPFLLRQGFYALGLLRRFDVGGRVGAHQPAFDQPGAILPERRQLARQRPVTQTALRQP					
	130	140	150	160	170	180
	190	200	210	220	230	240
g032	PQRRRKIAPRQVLRHAACIFRRHLCCQCKQFFQIAPVCRNRVLRLLALAHADVFIQISVKIRR					

g033.seq

1	ATGGCGGCGG	CGGACAAACT	CTTGGGCGGC	GACCGCCGCA	CGCTCGCCAT
51	CATCGGAGAC	GGCGCGATGA	CGGCGGGGCA	GGCGTTTGAA	GCCTTGAATT
101	GCGCGGCGCA	TATGGATGTG	GATTTGCTGT	TCGTCTCAA	CGACAACGAA
151	ATGTCGATT	CCCCCAACT	CGGCGCGTTG	CCCCAATATC	TTGCGACGAA
201	CGTCGTGCG	GCATATGCACG	GACTGTTGAG	TACCGTCAA	CGCAAAcgg
251	GCAAGGTATT	AGACAAAATA	CCCGGCGCGA	TGGagtTTGC	CCAAAAAGTC
301	GAACAcaaaA	TCAAAAC CCT	TGCCGAAGAA	GCCGAACACG	CCAAACAGTC
351	CGTGTCTGCT	TTTGA AATT	TCGGCTTCCG	CTACACCGGC	CCCGTGGACG
401	GACATACAGT	CGAGAACTCG	TGCGAGCTAT	TGAAAGACTT	GGCGAGCCGC
451	AAAGGCCCTC	AGTTGCTGCA	CGTCATCACC	AAA AAGGGA	ACGGCTACAA
501	ACTCGCCGAA	AACGACCCcg	tcaAATACCA	CGCCGTCGcc	aACCTGccta
551	AAGAAAGGCG	GGCGCCAATg	ccGTCGAAA	AAGAACCCAA	GCCCGCCgCc
601	aaaccgACCT	ATACCCAAGT	ATTGCGCAA	TGGCTGTGCG	ACCGGGCGGC
651	GGCAGATTCC	CGACTGGTTG	CGATTACCCC	CGCCATGCBC	GAGGGGACGC

213

```

701 GACTGGTGGG GTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT tacCTTTGCC GCGGTTTGG CGTGCGAAGG
801 CATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTGCCCGT TTTGTTTGCC
901 GTCGACCGTG CGGGCATCGT CGGCGCGGAC GGTCCGACCC ATGCCGGCTT
951 GTACGATTTG AGCTTCTTGC GCTGTGTGCC GAACATGATT GTTGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGCTGC TTTGACCTG CTATCAGGCG
1051 GATGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCC TtcatTGCCT TCGGCAGTAT GGTCGCCACC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTt
1251 cgtcaaacCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCAcy
1301 accGCATCGT TACCCTTGAA GAAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTCTTGG AAGTGTGGC GAAACACGGC ATCTGCAAA CCGTTTGTCT
1401 TTTGGGCGTT GCCGATACCG TAACCGAACA CGGCGATCCG AAAAACTTT
1451 TGGACGATT GGGTTTGAGT GCCGAAGCGG TGGAACGCCG GGTGCGCGAG
1501 TGCTGCGCG ACCGTGATGC GGCAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 116; ORF 033.ng>:

g033.pep

```

1 MAAADKLGG DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVLNDNE
51 MSISPNGAL PKYLASNVR DMHGLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTAE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLDLRSR
151 KGPQLLHVIT KKGNGYKLA E NDPVKYHAVA NLPKEGGAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLO RAYDQLVHDI ALQNLPLVLA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLLSCTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGK IIRREGEKTA FIAFGSMVAT
401 ALAVAELKNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKHG ICKPVLALLGV ADTVTEHGDG KKLDDDLGLS AEAVERRVRE
501 WLPDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 117>:

m033.seq

```

1 ATGGCGGCGG CAGACAACT CTTGGGCAGC GACCGCCGCA GCGTCGCCAT
51 CATCGGCGAC GCGCGATGA CCGCGGGGCA GCGCTTTGAA GCCTTGAATT
101 GCGCaG.CGA TATGGATGT GATTGCTrG TCGTCTCAA CGACAACGAA
151 ATGTCGATTT CCCCCAAGT CCGCGCGCTG CCGAAATACC TTGCCAGCAA
201 CGTCGTGCGC GATATGCACG GCCTGTTGAG TACCGTCAAA CGCGAAACGG
251 GCAAGGTATT AGACAAAATA CCCGGCGCGA TGGAGTTTGC CCAAAAAGTC
301 GAACACAAAA TCAAAACCTT TGCCGAAGAA GCCGAACACG CCAAACAGTC
351 GCTGTCTTTG TTTGAAAAT TCGGCTTCCG CTACACCGGC CCCGTGGACG
401 GACACAACGT CGAAAATCTG GTGGACGTAT TGAAAGACTT GCGCAGCCGC
451 AAAGGCCCTC AGTTGCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGACCCCG TCAAATACCA CGCCGTGCGC AACCTGCCTA
551 AAGAAAGCGC GCGGCAATG CCGTCTGAAA AAGAACCCTA GCCCGCGGCC
601 AAACCGACCT ATACCCAAGT GTTCGGCAAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
701 GCTTGGTTGA GTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCACGCCGT TACCTTTGCC GCGGTTTGG CTTGCGAAGG
801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCTACG
851 ACCAACTGGT GCACGACATC GCCCTGCAAA ACCTACCCGT TTTGTTTGCC
901 GTCGACCGCG CGGGCATCGT CGGCGCGGAC GGCCCGACCC ATGCCGGTCT
951 GTACGATTTG AGCTTTTTCG GCTGCGTGCC GAACATGATT GTCGCCGCGC
1001 CGAGCGATGA AAACGAATGC CGCCTGTTGC TTTGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCCGTCCG CTATCCGCGC GGCACGGGTA CGGGCGCGCC
1101 GGTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCCC
1201 GCATTGGCGG TTGCCGAAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCCTTGCC CGAAGCCACG

```

214

```

1301 ACCGCATCGT TACCCCTGAA GAAACGCCG AACAGGGCGG CGCAGGCGGC
1351 GCGGTGCTGG AAGTATTGGC GAAACACGGC ATCTGCAAAC CCGTTTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGAACGCGG TGTGCGCGG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 118; ORF 033>:

m033.pep

```

1 MAAADKLLGS DRRSVAIIGD GAMTAGQAFE ALNCAXDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTAAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLDLRSR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA
301 VDRAGIVGAD GPTHAGLYDL SFLRCVPNMI VAAPSDENEC RLLSTCYQA
351 DAPAAVRYPR GTGTGAPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401 ALAVAELKNA TVADMRVFKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGG
451 AVLEVLAKEG ICKPVLLLG V ADTVTGHGDP KKLLDDLGLS AEAVERRVRA
501 WLSDRDAAN*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 119>:

a033.seq

```

1 ATGGCGGCGG CGGACAAACA GTTGGGCAGC GACCGCCGCA GCGTCGCCAT
51 CATCGGCGAC GCGCGATGA CGGCGGGTCA GCGCTTTGAA GCCTTGAAC
101 GCGCGGGCGA TATGGATGTG GATTGCTGG TCGTCTCAA CGACAACGAA
151 ATGTGATTT CCCCCAAGT CGGTGCGTTG CCCAAATACC TTGCCAGCAA
201 CGTCGTGCGC GATATGCACG GACTGTTGAG TACCGTCAA GCGCAAACGG
251 GCAAGGTATT AGACAAAATA CCGGCGCGCA TGGAGTTGTC CCAAAAAGTC
301 GAACATAAAA TCAAACCCT TGCCGAAGAA GCCGAACAG CCAACAGTC
351 ACTGTCTTTG TTGAAAAC TCGGCTTCCG CTATACCGGC CCCGTGGACG
401 GACACAACGT CGAAAATCTG GTCGATGTAT TGAAGACCT GCGCGGACG
451 AAAGGCCCGC AGCTTCTGCA CGTCATCACC AAAAAGGGCA ACGGCTACAA
501 ACTCGCCGAA AACGATCCCG TCAAATACCA CGCCGTCGCC AACCTGCCTA
551 AAGAAAGCGC GCGCAAATG CCGTCTGAAA AAGAACCCTA GCCCGCGGCC
601 AAACCGACCT ATACCAAGT GTTCGGCAA TGGCTGTGCG ACCGGGCGGC
651 GGCAGATTCC CGACTGGTTG CGATTACCCC CGCCATGCGC GAGGGCAGCG
701 GCTTGGTTGA GTTTGAACAA CGATTCCCCG ACCGCTATTT CGATGTCGGC
751 ATCGCCGAGC AGCAGCCGT TACCTTTGCC GCGGTTTGG CTTGCGAAGG
801 GATGAAGCCC GTCGTGGCGA TTTATTCCAC CTTTTTACAA CGCGCTACG
851 ACCAACTGGT GACGACATC GCCCTGCAA ACCTGCCCGT TTTGTTGCC
901 GTCGACCGCG CGGCATCGT CGGCGCGGAC GGGCCGACCC ATGCCGTTT
951 GTACGATTGA AGCTTTTTCG GCTGCATTCC GAATATGATT GTCGCCGCGC
1001 CGAGCGATGA AAATGAATG CGCCTGCTGC TTTCGACCTG CTATCAGGCA
1051 GACGCGCCCG CCGCGTCCG CTATCCGCGC GGCACGGGTA CCGGCGTGCC
1101 GGTTTCAGAC GGCATGGAAA CCGTGGAAAT CGGCAAGGGC ATTATCCGCC
1151 GCGAAGGTGA GAAAACCGCA TTCATTGCCT TCGGCAGTAT GGTCGCCCTT
1201 GCATTGGCGG TCGCCGGAAA ACTGAACGCC ACCGTCGCCG ATATGCGCTT
1251 CGTCAAACCG ATAGACGAAG AGTTGATTGT CCGCTTTGCC CGAAGCCAG
1301 ACCGATCGT TACCCCTGAA GAAAACGCCG AACAGGGCGG CGCAGGCAGC
1351 GCGGTGCTGG AAGTGTGGC GAAACACGGC ATCTGCAAAC CCGTCTTGCT
1401 TTTGGGCGTT GCCGATACCG TAACCGGACA CGGCGATCCG AAAAACTTT
1451 TAGACGATTT GGGCTTGAGT GCCGAAGCGG TGAACGCGG TGTGCGCGG
1501 TGGCTGTCGG ATCGGGATGC GGCAAATTAA

```

This corresponds to the amino acid sequence <SEQ ID 120; ORF 033.a>:

a033.pep

```

1 MAAADKQLGS DRRSVAIIGD GAMTAGQAFE ALNCAGDMDV DLLVVLNDNE
51 MSISPNVGAL PKYLASNVR DMHLLSTVK AQTGKVLDKI PGAMEFAQKV
101 EHKIKTAAEE AEHAKQSLSL FENFGFRYTG PVDGHNVENL VDVLEDLRGR
151 KGPQLLHVIT KKGNGYKLAE NDPVKYHAVA NLPKESAAQM PSEKEPKPAA
201 KPTYTQVFGK WLCDRAAADS RLVAITPAMR EGSGLVEFEQ RFPDRYFDVG
251 IAEQHAVTFA GGLACEGMKP VVAIYSTFLQ RAYDQLVHDI ALQNLPLVFA

```

```

301 VDRAGIVGAD GPTHAGLYDL SFLRCIPNMI VAAPSDENEC RLLSTCYQA
351 DAPAAVRYPR GTGTGVPVSD GMETVEIGKG IIRREGEKTA FIAFGSMVAP
401 ALAVAGKUNA TVADMRFVKP IDEELIVRLA RSHDRIVTLE ENAEQGGAGS
451 AVLEVLAHKG ICKPVLLLG VADTVTGHGDP KKLLDDLGLS AEAVERRVRA
501 WLSDRDAAN*

```

m033/a033 98.4% identity over a 509 aa overlap

m033.pep	10	20	30	40	50	60
	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSISPNVGAL					
a033	MAAADKQLGSDRRSVAIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSISPNVGAL					
	10	20	30	40	50	60
m033.pep	70	80	90	100	110	120
	PKYLASNVRDMHGLLSTVKAQTGKVLDPGAMEFAQKVEHKIKTLAEAEHAKQSLSL					
a033	PKYLASNVRDMHGLLSTVKAQTGKVLDPGAMEFAQKVEHKIKTLAEAEHAKQSLSL					
	70	80	90	100	110	120
m033.pep	130	140	150	160	170	180
	FENFGFRYTGPDGHNVENLVDVLKDLRSRKGQQLLHVITKKNGYKLAENDPVKYHAVA					
a033	FENFGFRYTGPDGHNVENLVDVLEDLGRKGQQLLHVITKKNGYKLAENDPVKYHAVA					
	130	140	150	160	170	180
m033.pep	190	200	210	220	230	240
	NLPKESAAQMPSEKEPKPAKPTYTQVFGKWLCDRAAADSRVLAITPAMREGSGLVEFEQ					
a033	NLPKESAAQMPSEKEPKPAKPTYTQVFGKWLCDRAAADSRVLAITPAMREGSGLVEFEQ					
	190	200	210	220	230	240
m033.pep	250	260	270	280	290	300
	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLVLF					
a033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLVLF					
	250	260	270	280	290	300
m033.pep	310	320	330	340	350	360
	VDRAGIVGADGPTHAGLYDLSFLRCVPMIVAAPSDENECRLLSTCYQADAPAAVRYPR					
a033	VDRAGIVGADGPTHAGLYDLSFLRCIPNMIVAAPSDENECRLLSTCYQADAPAAVRYPR					
	310	320	330	340	350	360
m033.pep	370	380	390	400	410	420
	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFAFGSMVAPALAVAEKLNATVADMRFVKP					
a033	GTGTGVPVSDGMETVEIGKGIIRREGEKTAFAFGSMVAPALAVAGKLNATVADMRFVKP					
	370	380	390	400	410	420
m033.pep	430	440	450	460	470	480
	IDEELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAHKGICKPVLLLGVDVTVTGHGDP					
a033	IDEELIVRLARSHDRIVTLEENAEQGGAGSAGVLEVLAHKGICKPVLLLGVDVTVTGHGDP					
	430	440	450	460	470	480
m033.pep	490	500	510			
	KKLLDDLGLSAEAVERRVRAWLSDRDAANX					
a033	KKLLDDLGLSAEAVERRVRAWLSDRDAANX					
	490	500	510			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 033 shows 98.4% identity over a 509 aa overlap with a predicted ORF (ORF 033.ng) from *N. gonorrhoeae*:

m033/g033

m033.pep	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAXDMDVDLLVVLNDNEMSI SPNVGAL	60
g033	MAAADKLLGSDRRSVAIIGDGAMTAGQAFEALNCAGDMDVDLLVVLNDNEMSI SPNVGAL	60
m033.pep	PKYLASNVRDMHGLLSTVKAQTGKVLDKI PGAMEFAQKVEHKIKTLAEAEAHAKQSLSL	120
g033	PKYLASNVRDMHGLLSTVKAQTGKVLDKI PGAMEFAQKVEHKIKTLAEAEAHAKQSLSL	120
m033.pep	FENFGFRYTGPDGHNVENLVDVLKDLRSRKGQPLLHVITKKNGYKLAENDPVKYHAVA	180
g033	FENFGFRYTGPDGHNVENLVDVLKDLRSRKGQPLLHVITKKNGYKLAENDPVKYHAVA	180
m033.pep	NLPKESAAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVLAITPAMREGSGLVEFEQ	240
g033	NLPKEGGAQMPSEKEPKPAAKPTYTQVFGKWLCDRAAADSRVLAITPAMREGSGLVEFEQ	240
m033.pep	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA	300
g033	RFPDRYFDVGIAEQHAVTFAGGLACEGMKPVVAIYSTFLQRAYDQLVHDIALQNLPLVFA	300
m033.pep	VDRAGIVGADGPTHAGLYDLSFLRCVPMIVAAPSDENECRLLSTCYQADAPAAVRYPR	360
g033	VDRAGIVGADGPTHAGLYDLSFLRCVPMIVAAPSDENECRLLSTCYQADAPAAVRYPR	360
m033.pep	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVAPALAVA EKL NATVADMR FVKP	420
g033	GTGTGAPVSDGMETVEIGKGIIRREGEKTAFIAFGSMVATALAVA EKL NATVADMR FVKP	420
m033.pep	IDIELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVL LLGVADTVTGHGDP	480
g033	IDIELIVRLARSHDRIVTLEENAEQGGAGGAVLEVLAKHGICKPVL LLGVADTVTGHGDP	480
m033.pep	KKLLDDLGLSAEAVERRVRAWLSRDAANX	510
g033	KKLLDDLGLSAEAVERRVREWLPRDAANX	510

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 121>:

g034.seq

1	ATGAGCCGTT	TATGGTTTTT	TGCCGTAAAA	AACATTATAA	TCCGCCTTAT
51	TTACCTATTG	CCCAAGGAGA	CACAAATGGC	ACTCGTATCC	ATGCGCCAAC
101	TGCTTGACCA	CGCCGCCGAA	AACAGCTACG	GCCTGCCCCG	GTTCACCGTC
151	AACAACCTCG	AACAAATGCG	CGCCATTATG	GAAGCCGCCG	ACCAAGTCAA
201	CGCGCCCGTC	ATCGTACAGG	CGAGCGCAGG	TGCGCGCAAA	TACGcgGCG
251	CGCCGTTTTT	GCGCCACCTG	ATTCTGGCGG	CAGTCGAAGA	ATTTCCGCAC
301	ATCCCCGTCG	TGATGCACCA	AGACCACGGC	GCATCGCCCC	ACGTgtgCCA
351	ACGCTCCATC	CAACTGGGCT	TCTCCTCCGT	GATGATGGAC	GGCTCTTTGC
401	TCGAAGACGG	CAAAACCCCT	TCTTCTTACG	AATACAACGT	CAACGCCACC
451	CGTACCGTCG	TCAACTTCTC	CCACGCTGCG	GGCGTGTCGG	TCGAAGGCGA
501	AATCGGCGTA	TTGGGCAACC	TCGAAACCGG	CGAAGCAGGC	GAAGAAGACG
551	GAGTGGGCGC	GGCAGGCAAA	CTCTCACACG	ACCAAATGCT	CACCAGCGTT
601	GAAGATGCCG	TGCGTTTCGT	TAAAGATACC	GGCGTTGACG	CATTGGCGAT
651	TGCCGTCGGC	ACCAGCCACG	GCGCATACAA	ATTCACCCGT	CCGCCACAG

218

```

451 CGTACCGTGG TTAATTTCTC CCACGCCTGC GGCCTATCCG TTGAAGGCGA
501 AATCGGCGTA TTGGGCAACC TCGAAACTGG CGAAGCCGGC GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCCACG ACCAAATGCT CACCAGCGTC
601 GAAGATGCCG TCGGTTTCGT TAAAGATACC GGCCTTGACG CATGGCGAT
651 TGCCGTCGGC ACCAGCCACG GCGCGTACAA ATTCACCCGT CCGCCACAG
701 GCGACGTGTT GCGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGGCTCCAGC TCCGTTCCGC AAGAATGGCT
801 GAAAGTCATC AACGAATACG GCGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAC ACGGCGTGCG TAAAGTCAAC
901 ATCGATACCG ACTTGCGCCT TGCTTCCACC GGC GCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TCGATCCGCG CAAATATTG AGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GCTACCTCGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAT CAAACCGGT TCCTTGAAA AAATGGCAAA
1101 CCGTTATGCC AAGGCGAAT TGAACCAAT CGTCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 126; ORF 034.a>:

```

a034.pep
1 MSRLWFFAAK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
51 NNLEQMRAIM EAADQVNAPV IVQASAGARK YAGAPFLRHL ILAAVEEFPH
101 IPVVMHQDHG ASPDVCQRSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAVGK LSHDQMLTSV
201 EDAVRFKVDT GVDALAI AVG TSHGAYKFTR PPTGDLRLD RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKV NEYGGNIGET YGVPVEEIVE GIKHGVKRVN
301 IDTLRLAST GAVRRYLAEN PSDFDPRKYL SKTIEAMQI CLDRYLAFGC
351 EGQAGKIPV SLEKMANRYA KGELNQIVK*

```

m034/a034 96.9% identity over a 257 aa overlap

m034.pep	MSCLWFFAVKNIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM
a034	MSRLWFFAAKNIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM
m034.pep	EAADQVDAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI
a034	EAADQVNAPVIVQASAGARKYAGAPFLRHLILAAVEEFPHIPVVMHQDHGASPDVCQRSI
m034.pep	QLGFSSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLNLETGDAG
a034	QLGFSSVMMDGSLMEDGKTPSSYEYNVNATRTVVNFSHACGVSVEGEIGVLNLETGEAG
m034.pep	EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAI AVG TSHGAYKFTRPPTGDLRLD
a034	EEDGVGAVGKLSHDQMLTSVEDAVRFVKDTGVDALAI AVG TSHGAYKFTRPPTGDLRLD
m034.pep	RIKEIHQALPNTTHIVMH
a034	RIKEIHQALPNTTHIVMHGSSSVPQEWLKVINEYGGNIGETYGVPVEEIVEGIKHGVKRVN

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 034 shows 96.5% identity over a 257 aa overlap with a predicted ORF (ORF 034.ng) from *N. gonorrhoeae*:

m034/g034

m034.pep	MSCLWFFAVKNIIRLIYLLPKETQMALVSMRQLLDHAAEXSYGLPAFNVNNLXQMRAIM	60
g034	MSRLWFFAVKNIIRLIYLLPKETQMALVSMRQLLDHAAENSYGLPAFNVNNLEQMRAIM	60
m034.pep	EAADQVDAPVIVQASAGARKYAGAPFLRHLILA AVEVFPHIPVVMHQDHGASPDVCQRSI	120
g034	EAADQVNAPVIVQASAGARKYAGAPFLRHLILA AVEEFPHIPVVMHQDHGASPDVCQRSI	120
m034.pep	QLGFSSVMDGSLMEDGKTPSSYEYNVNATRTVVNF SHACGVSVEGEIGVLGNLETGDAG	180
g034	QLGFSSVMDGSLLEDGKTPSSYEYNVNATRTVVNF SHACGVSVEGEIGVLGNLETGEAG	180
m034.pep	EEDGVGAVGKLSHDQMLTSVEDAVCFVKDTGVDALAI AVGTSHGAYKFRPPTGDVLRID	240
g034	EEDGVGAAGKLSHDQMLTSVEDAVRFVKDTGVDALAI AVGTSHGAYKFRPPTGDVLRID	240
m034.pep	RIKEIHQALPNTHIVMH	257
g034	RIKEIHQALPNTHIVMHGSSSVPOEWLKVINEYGGNIGET YGVPVEEIVEGIKHGVRKVN	300

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 127>:

g036.seq

```

1  ATGCTGAAGC CGTGTGTTGGT ATACAGTGCC TGTGCGGCGG cgttgccTGC
51  GCGGACTTCG AGCAGCAGGC GTTGCGTGCC TTCGGGCAGA TGTGCGTACC
101 AATATTCGAG CAGGGCGGAC GCAACGCCCC GTCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTCTC GCCAAACGAT
201 AAAGGCGGCA ATCCTGCCGT CTTTTCCTCGC AAGGAAAACC TGTTCGGACG
251 GCGAAACAAG CGCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGTAT CGAGCGCGGC CAGTGCAGCG CAGTCCGACG GTGAGGCTGG
351 GCGGATGTTT ATGTTCGTGC CTTCCGTTCC GCCTGTTCTT TGGCAGTCAG
401 GCGGATTTTG TTGCGGACGT AGAGCAGTTC GCGGTGTGCC GCGCCAGTTG
451 CCGGATAGCC GCCGCCGAGG GCGAGCGCGA GAAAATCGGC GGCGGTCCGC
501 ATATCGGGTT TGCTGAGAA GGGCGGACGG TTTTCCAGTG CGAACGCACT
551 GCGGATGCGG TCTGAAAAGA CGTACCCCTC GGGGAGGCGA ATGTCTGCCG
601 CCCTACCGAC TTGATAATCG CTCAAACGGC GGCGGTTTCA CGTGTGGAAC
651 CACGCATAAA AACTTCGCC CATAACGCGC TCCGACGCGG CGAGTATGCA
701 GCTTTGCGGC GGCAGCAGCG AGGCGGCGGC ATCGAGCGTG GGGATGCCGA
751 TTAAAGCGGT GTCGAACGGC GTTGCCAAAC CTGCGCCAC CCGATGCCG
801 ATACGCAGTC CGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 128; ORF 036.ng>:

g036.pep

```

1  MLKPCLVYSA CAAALPARTS SSRRCVPSGR CAYQYSSRAD ATPRRRHSGA
51  VAIRCSSDSS GRFCQTIKAA ILPSFSARKT CSDGETSADS NWRCVHADGL
101 QTVSSAASAA QSDGEAGRMF MFVPSVPPVL WQSGRFCCGR RAVRRVPRQL
151 RDSRRRRRAR ENRRRSAYRV CLRRADGFPV RTHCRCRLKR RTPRGQCCLP
201 PYRLDNRSNG GGSACRTTHK TLRPYARPQR RVCSEAAAAA RRRHRAWGCR
251 LKACRTALPN LAPRRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 129>:

m036.seq

```

1  ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTGCTGTC
51  ACGGACTTCG AGCAGCAGGC GTTGCGTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GGCGGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTCTC GCCAAACGAT
201 AAAGGCGGCA ATCCG.CGT CTTTTCCTCGC AAGGAAAACC TGTTCGGACG
251 GCGAAACCAG TGCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301 CAGACGGCAT CGAGTGCAGC CAGCTCCTCA CAATCGGCAC AAACGGCAGC

```


220

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTCT GCCTGTTCTT TGGCAGTCAG
401 GGC GATTTT TCGCGACGT AGAGCAAACC GCGGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA AACTTTCGCC CATAAGAGCG TCCGTAGCGG CAAGGATGCA
701 GCTTTGCGGC GGC GCGAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGGGGT ATCAAACGGC GTTGCCAAAC CCTGAGCTAC ACCGATGCCG
801 ATACGCAGTC CGGTAA

```

This corresponds to the amino acid sequence <SEQ ID 130; ORF 036>:

```

m036.pep
  1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51 VAIRCSSDSS GRFCQTIKAA IPXSFSAARKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARRMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRPWLPMR ESRRQSAYPV CLRTAELLPA RTRCLCRLKR RIPPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*R QGCSFAAAAA RRRHRARVRR
251 LRGYQTALPN PELHRCRYAV R*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 131>:

```

a036.seq
  1 ATGCTGAAGC CGTGCGCCGT GTACAGTGCC TGTGCGGCGG TGTGCTGCTG
 51 ACGGACTTCG AGCAGCAGGC GTTGCGTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GCGCGCGGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCCGCCGT CTTTTCCGCG AAGGAAAACC TGTTCGGACG
251 GCGAAACCAG TCGGACTCA AATTGGCGTT GCGTCCACGC GGACGGGTTG
301 CAGACGGCAT CGAGCGCGGC GAGTGC GCGC CAATCGGCAT AAACGGCGCG
351 GCGGATGTTT ACAGGCGCGC CCTCCGTTCC GCCTGTTCTT TGGCAGTCAA
401 GGC GATTTT TCGCGACGT AGAGCAGCTC GCGGTGTGCC GCAGCGACGG
451 CGGGAAAACC GCCTTCAGCC GCCAGATTGA GGAAGTCGGC GCGGTCGGC
501 ATATCGGGTT TGCCTGAGAA GGGCGGACGG TTTTCCAGCG CGAACGCATT
551 GCCGATGCCG TCTGAAAAGG CGCATCCTTC CGGCAGCCGG ATGTCTGCCG
601 CCCGACCGAC CTGATAATCG CTCAAACGGC GCGGTTTCAG CGTGTGCAAC
651 CATGCATAAA AACTTTCGCC CATACTGCGC TCCGAGCGCG CAAGGATGCA
701 GCTTTGCGGC GGC GCGAGCG AGGCGGCGGC ATCGAGCGAG GGTACGCCGA
751 TTAAGGAGT ATCAAACGGC GTTGCCAAAC CTTGCGCCAC GCCGATGCCG
801 ATACGCAGTC CCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 132; ORF 036.a>:

```

a036.pep
  1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51 VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
101 QTASSAASAA QSA*TARRMF TGAPSVPPVL WQSRRFCCGR RAARRVPQRR
151 RENRLQPPD* GSRRRSAYRV CLRRADGFPA RTHCRCLKR RILPAAGCLP
201 PDRPDNRSNG GGSACRTMHK TLRPVVRPQR QGCSFAAAAA RRRHRARVRR
251 LKEYQTALPN LAPRRCRYAV P*

```

m036/a036 85.6% identity over a 270 aa overlap

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAI PWRRHSGAVAI RCSSDSS					
a036	MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAI PWRRHSGAVAI RCSSDSS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m036.pep	GRFCQTIKAAIPXSFSAARKT CSDGETSADSNWRCVHADGLQTASSAASSSQSAQTARRMF					
a036	GRFCQTIKAAIPPSFSARKT CSDGETSADSNWRCVHADGLQTASSAASAAQSAXTARRMF					
	70	80	90	100	110	120

221

	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMPRESRRQSAYPVCRLRTAELLPA					
a036	TGAPSVPPVLWQSGRFCCGRRRAARRVPQRRRENRLQPPDXGSRRRSAYRVCLRRADGFPA					
	130	140	150	160	170	180
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPKRQGCSFAAAAA					
a036	RTHCRCLKRRILPAAGCLPPDRPDNRSNGGGSACRTMHKTLRPYVRPQRQGCSFAAAAA					
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPKRQGCSFAAAAA					
a036	RTHCRCLKRRILPAAGCLPPDRPDNRSNGGGSACRTMHKTLRPYVRPQRQGCSFAAAAA					
	190	200	210	220	230	240
m036.pep	RRRHRARVRLRGYQTALPNPELHRCRYAVRX					
a036	RRRHRARVRLKEYQTALPNLAPRRCRYAVPX					
	250	260	270			
m036.pep	RRRHRARVRLRGYQTALPNPELHRCRYAVRX					
a036	RRRHRARVRLKEYQTALPNLAPRRCRYAVPX					
	250	260	270			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 036 shows 74.9% identity over a 271 aa overlap with a predicted ORF (ORF 036.ng) from *N. gonorrhoeae*:

m036/g036

	10	20	30	40	50	60
m036.pep	MLKPCAVYSACAAVLPARTSSSRRCVSSGRCVNQYSSRADAIWRRHSGAVAIRCSSDSS					
g036	MLKPCLVYSACAAALPARTSSSRRCVPSGRCAQYSSRADATPRRRHSGAVAIRCSSDSS					
	10	20	30	40	50	60
m036.pep	GRFCQTIKAAIPXSFSARKTCSGGETSADSNWRCVHADGLQTASSAASSSSQSAQTARRMF					
g036	GRFCQTIKAAILPSFSARKTCSGGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF					
	70	80	90	100	110	120
m036.pep	GRFCQTIKAAIPXSFSARKTCSGGETSADSNWRCVHADGLQTASSAASSSSQSAQTARRMF					
g036	GRFCQTIKAAILPSFSARKTCSGGETSADSNWRCVHADGLQTVSSAASAAQSDGEAGRMF					
	70	80	90	100	110	120
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMPRESRRQSAYPVCRLRTAELLPA					
g036	MFVPSVPPVLWQSGRFCCGRRAVRRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFPA					
	130	140	150	160	170	180
m036.pep	TGALSVRPVLWQSGRFCCGRRANRRVRHGRQDNRPWLPMPRESRRQSAYPVCRLRTAELLPA					
g036	MFVPSVPPVLWQSGRFCCGRRAVRRVPRQLRDSRRRGRARENRRRSAYRVCLRRADGFPA					
	130	140	150	160	170	180
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPKRQGCSFAAAAA					
g036	RTHCRCLKRRTPRGGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPPQRRVCSFAAAAA					
	190	200	210	220	230	240
m036.pep	RTRCLCRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPKRQGCSFAAAAA					
g036	RTHCRCLKRRTPRGGQCLPPYRLDNRSNGGGSACRTTHKTLRPYARPPQRRVCSFAAAAA					
	190	200	210	220	230	240
m036.pep	RRRHRARVRLRGYQTALPNPELHRCRYAVRX					
g036	RRRHRAGCRLKACRTALPNLAPRRCRYAVRX					
	250	260	270			
m036.pep	RRRHRARVRLRGYQTALPNPELHRCRYAVRX					
g036	RRRHRAGCRLKACRTALPNLAPRRCRYAVRX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 133>:

m036-1.seq

```

1  ATGCTGAAGC CGTGC GCCGT GTACAGTGCC TGTGCGGCGG TGTGCCTGC
51  ACGGACTTCG AGCAGCAGGC GTTGC GTGTC TTCGGGCAGA TGTGTGAACC
101 AATATTCGAG CAGGGCGGAC GCAATTCCTT GCGGCGCGCA TTCGGGCGCG
151 GTGGCAATCA GGTGCAGTTC GGATTCGTCG GGCAGGTTCT GCCAAACGAT
201 AAAGGCGGCA ATCCCGCCGT CTTTTCCTGC AAGGAAAACC TGTTCGGACG
251 CGGAAACCCAG TCGGACTCA AATTGGCGTT GCGTCCATGC GGACGGGTTG
301 CAGACGGCAT CGAGTGC GCGC CAGCTCCTCA CAATCGGCAC AAACGGCAGC

```

```

351 GCGGATGTTT ACGGGCGCGC TCTCCGTTTC GCCTGTTCTT TGGCAGTCAG
401 GGC GATTTTGG TTGCGGACGT AGAGCAAACC GCGGTGTGCG GCATGGACGG
451 CAGGATAACC GCCCTTGGCT GCCAATGCGA GAAAGTCGGC GGCAGTCGGC
501 ATATCCGGTC TGCCTGAGAA CGGCGGAGCT TCTTCCAGCG CGAACGCGCT
551 GCCTATGCCG TCTGAAAAGG CGCATCCCTC CGGCAGCCGG ATGTCTGCCG
601 CCCGCCCGAC CTGATAATCG CTCAAACGGT GGCAGTTCAG CGTATCGAAC
651 CATGCATAAA AACTTTCGCC CATACGAGCG TCCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 134; ORF 0036-1>:

```

m036-1.pep
  1 MLKPCAVYSA CAAVLPARTS SSRRCVSSGR CVNQYSSRAD AIPWRRHSGA
 51 VAIRCSSDSS GRFCQTIKAA IPPSFSARKT CSDGETSADS NWRCVHADGL
101 QTASSAASSS QSAQTARMF TGALSVRPVL WQSGRFCCGR RANRRVRHGR
151 QDNRFPWPMR ESRRSAYPV CLRTAELLPA RTRCLRLKR RIPPAAGCLP
201 PARPDNRSNG GSSAYRTMHK TLRPYERP*

```

m036-1/g036 76.8% identity in 228 aa overlap

m036-1.pep	10	20	30	40	50	60
	MLKPCAVYSACAAVLPARTSSSRRCVSSGRVCNQYSSRADAI P WRRHSGAVAI RCSSDSS					
g036	MLKPCLVYSACAAALPARTSSSRRCVPSGRCA YQYSSRADATPRRRHSGAVAI RCSSDSS					
	10	20	30	40	50	60
m036-1.pep	70	80	90	100	110	120
	GRFCQTIKAAIPPSFSARKT CSDGETSADS NWRCVHADGLQTASSAASSS QSAQTARMF					
g036	GRFCQTIKAAILPSFSARKT CSDGETSADS NWRCVHADGLQTVSSAASAAQSDGEAGRMF					
	70	80	90	100	110	120
m036-1.pep	130	140	150	160	170	180
	TGALSVRPVLWQSGRFCCGR RANRRVRHGRQDNR P WPMRESRRQSAYPVCLRTAELLPA					
g036	MFVPSVPPVLWQSGRFCCGRRAVRVPRLRDSRRRGRARENRRRSAYRVCLRRADGFPV					
	130	140	150	160	170	180
m036-1.pep	190	200	210	220	229	
	RTRCLRLKRRIPPAAGCLPPARPDNRSNGGSSAYRTMHKTLRPYERPX					
g036	RTHCRCLKRRTPRGGQCLPPYRLDNRSNGGSACRTTHKTLRPYARPQRRVCSFAAAAA					
	190	200	210	220	230	240
g036	RRRHRAWGCRLKACRTALPNLAPRRCRYAVRX					
	250	260	270			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 135>:

```

g038.seq
  1 ATGACTGATT TCCGCAAGA TTTCTCAAA TTCTCCCTCG CCCAAAATGT
 51 TTTGAAATTC GCGGAATTTA CCACCAAAGC CGGACGGCGG TCGCCCTATT
101 TCTTCAATGC CGGCTCTTTC AACGACGGCG CGTCCACGCT GCAACTGGCA
151 AAATTCTATG CACAATCCAT CATTGAAAGC GGCAATCCGAT TCGATATGCT
201 GTTCGGCCCC GCCTACAAAG GCATTATTTT GCGGGCGGCA ACCGCGATGA
251 TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301 GCCAAAGACC GCGGCGAAGG CGGCGTGTTC GTCGGCGCGC CGCTTAAAGG
351 GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401 AATCAATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGCG CGGTGTCGCC
451 ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTAAATTGT CCGCCGTTCA
501 GGAAGTGGAA AAACAATACG GCCTGCCCGT CGCCCCATC GCCAGCCTGA
551 ACGATTGTGT TATCCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
601 GAACCCGTCC GCACCTACCG CCGGCAGTAC GCGGTAGAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 136; ORF 038.ng>:

```
g038.pep
  1  MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGASTLQLA
 51  KFYAQSIIES GIRFDMLEGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101  AKDRGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
151  IALDRMEKGT GKLSAVQVEV KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201  EPVRTYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 137>:

```
m038.seq
  1  ATGACCGATT TCCGCCAAGA TTTCTCCTCAA TTCTCCCTCG CCCAAAATGT
 51  TTTGAAATTC GGCGAATTTC CCACCAAGGC AGGACGGCGG TCGCCCTATT
101  TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
151  AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201  GTTCGGTCCC GCCTACAAAG GCATTATTTT GCGCGCGGCA ACCGCGATGA
251  TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301  GCCAAAGACC ACGGCGAAGG CGGCGTGTG GTCGGCGCGC CGCTTAAAGG
351  GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401  AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGC CGGTGTGCGC
451  ATCGCGCTCG ATCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
501  GGAAGTGGAA AAACAATACG GkCTGCCCCT CGCCCCATC GCCAGCCTGA
551  ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
601  GAACCCGTCC GAGCCTACCG TCGGCAGTAC GCGGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 138; ORF 038>:

```
m038.pep
  1  MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
 51  KFYAQSIIES GIRFDMLEGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101  AKDHGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
151  IALDRMEKGT GELSAVQVEV KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201  EPVRAYRRQY GVE*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 139>:

```
a038.seq
  1  ATGACCGATT TCCGCCAAGA TTTCTCCTCAA TTCTCCCTCG CCCAAAATGT
 51  TTTGAAATTC GGCGAATTTC CCACCAAGGC CGGACGGCGG TCGCCCTATT
101  TCTTCAATGC CGGCCTCTTT AACGACGGCT TGTCCACGCT GCAACTGGCA
151  AAATTTTACG CACAATCCAT CATTGAAAGC GGCATCCGAT TCGATATGCT
201  GTTCGGCCCC GCCTACAAAG GCATTATTTT GCGCGCGGCA ACCGCGATGA
251  TGCTGGCGGA AAAAGGCGTG AACGTCCCGT TTGCCTACAA CCGCAAAGAA
301  GCCAAAGACC ACGGCGAAGG CGGCGTGTG GTCGGCGCGC CGCTTAAAGG
351  GCGCGTGCTG ATTATCGACG ACGTGATTTC CGCCGGCACA TCCGTACGCG
401  AATCGATCAA ACTGATTGAA GCGGAGGGTG CAACCCCGC CGGTGTGCGC
451  ATCGCGCTCG ACCGCATGGA AAAAGGCACG GGTGAATTGA GCGCGGTTCA
501  GGAAGTGGAA AAACAATACG GCCTGCCCCT CGCCCCATC GCCAGCCTGA
551  ACGATTTGTT TATTCTGTTG CAAAACAACC CCGAATTCGG ACAGTTCCTC
601  GAACCCGTCC GAGCCTACCG TCGGCAGTAC GCGGTAGAAT AA
```

This corresponds to the amino acid sequence <SEQ ID 140; ORF 038.a>:

```
a038.pep
  1  MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA
 51  KFYAQSIIES GIRFDMLEGP AYKGIILAAA TAMMLAEKGV NVPFAYNRKE
101  AKDHGEGGVL VGAPLKGRVL IIDDVISAGT SVRESIKLIE AEGATPAGVA
151  IALDRMEKGT GELSAVQVEV KQYGLPVAPI ASLNDLFILL QNNPEFGQFL
201  EPVRAYRRQY GVE*
```

m038/a038 100.0% identity over a 213 aa overlap

```

          10      20      30      40      50      60
m038.pep  MTDFRQDFLK FSLAQNVLFK GEFTTKAGRR SPYFFNAGLF NDGLSTLQLA KFYAQSIIES
```

224

```

|||||
a038      MTDFRQDFLKFSLAQNVLKFGFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
           10      20      30      40      50      60

           70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
           |||||
a038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
           70      80      90      100     110     120

           130     140     150     160     170     180
m038.pep  IIDDISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
           |||||
a038      IIDDISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
           130     140     150     160     170     180

           190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
           |||||
a038      ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
           190     200     210

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 038 shows 98.1% identity over a 213 aa overlap with a predicted ORF (ORF 038.ng) from *N. gonorrhoeae*:

m038/g038

```

           10      20      30      40      50      60
m038.pep  MTDFRQDFLKFSLAQNVLKFGFTTKAGRRSPYFFNAGLFNDGLSTLQLAKFYAQSIIES
           |||||
g038      MTDFRQDFLKFSLAQNVLKFGFTTKAGRRSPYFFNAGLFNDGASTLQLAKFYAQSIIES
           10      20      30      40      50      60

           70      80      90      100     110     120
m038.pep  GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDHGEGGVLVGAPLKGRVL
           |||||
g038      GIRFDMFLFGPAYKGIILAAATAMMLAEKGVNVPFAYNRKEAKDRGEGGVLVGAPLKGRVL
           70      80      90      100     110     120

           130     140     150     160     170     180
m038.pep  IIDDISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGELSAVQEVEKQYGLPVAPI
           |||||
g038      IIDDISAGTSVRESIKLIEAEGATPAGVAIALDRMEKGTGKLSAVQEVEKQYGLPVAPI
           130     140     150     160     170     180

           190     200     210
m038.pep  ASLNDLFILLQNNPEFGQFLEPVRAYRRQYGVEX
           |||||
g038      ASLNDLFILLQNNPEFGQFLEPVRTYRRQYGVEX
           190     200     210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 141>:

```

g039.seq
1   ATGCCGTCCG AACCACCTGC CGCTTCAGAC GGCATCAAAC CGACACACAC
51  CGAGAAAACA TCATGCCCGC CTGTTTCTGT CCGCACTGCA AAACCCGCCT
101 CTGGGTCAAA GAAAcccagC TCAAcgtCgC ccaagGCTTC GTCGTCTgcc
151 aaAAAtgcga agGGCTgttt aaAgccaaaG accAtctggc aaGcacGAAA
201 gaacctatat tcaacgattg gcccgaagct gtttcgggat gTcaaaCTCG

```

225

```

251 TCcaccgcaT cggcacgcac gccattagca aGAaacagat gtccccgcgac
301 gaaatCgccg atatcctcaa cggcggtaca acCCTGCACG ATACGCCGCC
351 CGCAACCGCC GCTGCCGCac ctGCCGCCGC ACCGCaggTT TCCGTACCGC
401 CCGCCCGTCA GGAAGGGCTC AACTGGACTA TTGCAACCCT GTTCGCACTT
451 ATCGTCCTCA TTATGCAGCT TTCTACCTC TTCATCCTAT GA

```

This corresponds to the amino acid sequence <SEQ ID 142; ORF 039.ng>:

```

g039.pep
1  MPSEPPAASD GIKPTHTEKT SCPPVSVRTA KPASGSKKPS STSPKASSSA
51  KNAKGCLKPK TIWQARKNLY STIGPKLFRD VKLVHRIGTH AISKQMSRD
101 EIADILNGGT TLHDTTPATA AAAPAAPOV SVPPARQEG L NWTIATLFLAL
151 IVLIMQLSYL FIL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 143>:

```

m039.seq
1  ATGCCGTCCG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51  CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
101 CTGGGTCAAA GAAACCCAAC TCAATGTCGC CGnnnnnnnnn nnnnnnnnnn
151 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
201 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnCCC GAGGCTGT TT
251 CGGATGTCAA ACTCGTTCAC CGTATCGGCA CGCGCGCCAT CGGCAAGAAA
301 CAGATTTCCC GTGACGAAAT CGCCGGCATC CTCAACGGCG GTACAACCCA
351 GCCCGATATT CCGCCGCAA CCGCCGCCAC CCCTGCTGCC GCACCGCAGG
401 TTACCGTACC GCCCGCGCG CCGCCCGTC AGGATGGGTT CAACTGGACG
451 ATTGCAACCC TGTTCGCCCT TATCGTCTC ATTATGCAGC TTCCTACCT
501 CGTCATCCTA TGA

```

This corresponds to the amino acid sequence <SEQ ID 144; ORF 039>:

```

m039.pep
1  MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPPXXXXXX
51  XXXXXXXXXX XXXXXXXXXX XXXXXXXXXP EAVSDVKLVH RIGTRAIGKK
101 QISRDEIAGI LGGTTQPD I PPATAATPAA APQVTVPAA PARQDGFNWT
151 IATLFLALIVL IMQLSYLVIL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 145>:

```

a039.seq
1  ATGCCGTCTG AACCGCCTTA CGCCTCAGAC GGCATCAAAC CTGACACACA
51  CGAGGAAATA CCATGCCCGC CTGTTTCTGC CCCCACTGCA AAACCCGTCT
101 CTGGGTCAAA GAAACCCAAC TCAATGTCGC CCAAGGCTTC GTGCTCTGCC
151 AAAAATGCGA AGGAATGTTT AAAGCCAAAG ACCATCTGGC AAGCACGAAA
201 GAACCCATAT TCAACGATT. TGCCCGAAGC TGTTTCGGAT GTCAAACCTG
251 TTCACCGCAT CGGCACGAGC GCCATCGGCA AGAAACAGAT TTCCCGTGAC
301 GAAATCGCCG GCATCCTCAA CGGCGGCACA ACCAGCCCG ATATTCCGCC
351 CGCAACCGCC GCCACCCCTG CTGCCGCACC GCAGGTTACC GTACCGCCCG
401 CCGCGCCCGC CCGTCAGGAT GGGTTCAACT GGACGATTGC AACCTGTTT
451 GCCCTTATCG TCCTCATTAT GCAGCTTTC TACCTCGTCA TCCTATGA

```

This corresponds to the amino acid sequence <SEQ ID 146; ORF 039.a>:

```

a039.pep
1  MPSEPPYASD GIKPDTHEEI PCPPVSAPTA KPVSGSKKPN SMSPKASSSA
51  KNAKECLKPK TIWQARKNPY STIXPEAVSD VKLVHRIGTS AIGKKQISRD
101 EIAGILNGGT TQPDIPPATA ATPAAAPQVT VPPAAPARQD GFNWTIATLE
151 ALIVLIMQLS YLVIL*

```

m039/a039 79.4% identity over a 170 aa overlap

```

          10      20      30      40      50      60
m039.pep  MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPN SMSPPXXXXXXXXXXXXXXXXX
          |||
a039      MPSEPPYASDGIKPDTHEEIPCPPVSAPTAKPVSGSKKPN SMSPKASSSAKNAKECLKPK

```

```

701 GCGACGTATT GCGTATCGAC CGCATCAAGG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA CGgctCCAGC TCCGTTCCGC AAGAatgGCT
801 GAAAGTCATC AACGAATACG GCGGCAATAT CGGCGAAACC TACGGCGTGC
851 CGGTTGAAGA AATCGTCGAA GGCATCAAAC ACGGCGTGCG CAAAGTCAAC
901 ATCGATACCG ACCTGCGCCT CGCTTCCACC GCGCGGTAC GCCGCTACCT
951 TGCCGAAAAC CCGTCCGACT TTGATCCGCG CAAATACTTG GGCAAAACCA
1001 TTGAAGCGAT GAAGCAAATC TGCCTCGACC GTTATCTTGC GTTCGGTTGC
1051 GAAGGTCAGG CAGGCAAAT CAAACCTGTT TCGTTGGAAA AAATGGCAAG
1101 CCGTTATGCC AAGGCGCAAT TGAACCAAT CGTCAAATAA

```

This corresponds to the amino acid sequence <SEQ ID 122; ORF 034.ng>:

g034.pep

```

1 MSRLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE NSYGLPAFNV
51 NNLEQMRAIM EAADQVNPV IVQASAGARK YAGAPFLRHL ILAAVEEFPF
101 IPVVMHODHG ASPDVCQRSI QLGFSVMMD GSLLDGKTP SSYEYNVNAT
151 RTVVNFSHAC GVSVEGEIGV LGNLETGEAG EEDGVGAAGK LSHDQMLTSV
201 EDAVRFVKDT GVDALAIAGV TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMHGSS SVPQEWLKI NEYGGNIGET YGVFVEEIVE GIKHGVKRVN
301 IDTDLRLAST GAVRRYLAEN PSDFDPRKYL GKTIEAMKQI CLDRYLAFGC
351 EGQAGKIKPV SLEKMASRYA KGELNQIVK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 123>:

m034.seq (partial)

```

1 ATGAGCTGTT TATGGTTTTT TGCTGTAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA WACAGCTACG GCyTGCCGGC GTTCAACGTC
151 AACAACTCG WACAGATGCG CGCCATCATG GAGGCTGCAG ACCAAGTCGA
201 CGCCCCGTC ATCGTACAGG CGAGTGCCGG TGCGCGCAA TATCGGGTG
251 CGCCGTTTTT ACGCCACCTG ATTTTGGCGG CTGTCGAAGT ATTTCCACAC
301 ATCCCCGTCG TCATGCACCA AGACCACGGC GCATCACCCG ACGTGTGCCA
351 ACGCTCCATC CAACTGGGCT TCTCCTCTGT AATGATGGAC GGCTCGCTGA
401 TGAAGACGG CAAAACCCCT TCTTCTTACG AATACAACGT CAACGCCACA
451 CGTACCGTGG TTAACCTTCTC CCACGCTTGC GCGGTATCCG TTGAAGCGGA
501 AATCGGCGTA TTGGGCAACC TCGAAACCGG CGATGCAGGC GAAGAAGACG
551 GTGTAGGCGC AGTGGGCAAA CTTTCCACG ACCAAATGCT GACCAGCGTC
601 GAAGATGCCG TATGTTTCGT TAAAGATACC GCGTTGACG CATTGGCTAT
651 TGCCGTCGGC ACCAGCCACG GCGCATACAA ATTCACCCGT CCGCCACAG
701 GCGATGTATT ACGTATCGAC CGCATCAAAG AAATCCACCA AGCCCTGCCC
751 AATACACACA TCGTGATGCA C...

```

This corresponds to the amino acid sequence <SEQ ID 124; ORF 034>:

m034.pep (partial)

```

1 MSCLWFFAVK NIIIRLIYLL PKETQMALVS MRQLLDHAAE XSYGLPAFNV
51 NNLXQMRAIM EAADQVDAPV IVQASAGARK YAGAPFLRHL ILAAVEVFPF
101 IPVVMHODHG ASPDVCQRSI QLGFSVMMD GSLMEDGKTP SSYEYNVNAT
151 RTVVNFSHAC GVSVEGEIGV LGNLETGDAG EEDGVGAVGK LSHDQMLTSV
201 EDAVCFVKDT GVDALAIAGV TSHGAYKFTR PPTGDVLRID RIKEIHQALP
251 NTHIVMH...

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 125>:

a034.seq

```

1 ATGAGCCGTT TATGGTTTTT TGCCGCAAAA AACATTATAA TCCGCCTTAT
51 TTACCTATTG CCCAAGGAGA CACAAATGGC ACTCGTATCC ATGCGCCAAC
101 TGCTTGATCA TGCTGCCGAA AACAGCTACG GCCTGCCCGC GTTCAACGTC
151 AACAACTCG AACAAATGCG CGCCATTATG GAAGCGCGCC ACCAAGTCAA
201 CGCGCCCGTC ATCGTACAGG CGAGCGCAGG TGCGCGCAA TACGCGGGCG
251 CGCCGTTTTT GCGCCACCTG ATTTTGGCGG CTGTCGAAGA ATTTCCGCAC
301 ATCCCCGTCG TGATGCACCA AGACCACGGC GCATCGCCCG ACGTGTGCCA
351 ACGTCCATC CAACTGGGCT TTTCTCTCGT GATGATGGAC GGCTCGCTGA
401 TGAAGACGG CAAAACCCCT TCTTCTTATG AATACAACGT CAACGCCACC

```

226

	10	20	30	40	50	60
m039.pep	70	80	90	100	110	120
	XXXXXXXXXXXXXXXXXXXX	PEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPDI				
a039	:	:				
	70	80	90	100	110	
	TIWQARKNPYSTIX----	PEAVSDVKLVHRIGTSAIGKKQISRDEIAGILNGGTTQPDI				
	70	80	90	100	110	
m039.pep	130	140	150	160	170	
	PPATAATPAAAPQVTVP	PAAAPARQDGFNWTIATL	FALIVLIMQLSYLVILX			
a039						
	120	130	140	150	160	
	PPATAATPAAAPQVTVP	PAAAPARQDGFNWTIATL	FALIVLIMQLSYLVILX			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 039 shows 60.8% identity over a 171 aa overlap with a predicted ORF (ORF 039.ng) from *N. gonorrhoeae*:

m039/g039

	10	20	30	40	50	60
m039.pep	MPSEPPYASDGIKPD	THEEIPCPVSAPTAKPV	SGSKKPNSMSPXXXXXX			
g039						
	10	20	30	40	50	60
	MPSEPPAASDGIKPT	HEKTSCTPPVSVRTAK	PASGSKKPSSTSPKASS	SAKNAKGCLKPK		
	10	20	30	40	50	60
m039.pep	70	80	90	100	110	120
	XXXXXXXXXXXXXXXXXX	PEAVSDVKLVHRIGTRAIGKKQISRDEIAGILNGGTTQPDI				
g039	:	:				
	70	80	90	100	110	
	TIWQARKNLYSTIG----	PKLFRDVKLVHRIGTHAISKQMSRDEIADILNGGTTLHDT				
	70	80	90	100	110	
m039.pep	130	140	150	160	170	
	PPATAAT-PAAAPQVTVP	PAAAPARQDGFNWTIATL	FALIVLIMQLSYLVILX			
g039						
	120	130	140	150	160	
	PPATAAAAPAAAPQVSV	PPA--RQEGLNWTIATL	FALIVLIMQLSYLFILX			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 147>:

g040.seq

1	ATGAACGCGC	CGACAGCTT	TGTCGCCAC	TTCCGCGAAG	CCGCCCCCTA
51	CATCCGCCAA	ATGCGCGGCA	CGACACTGGT	CGCCGGCATA	GAaggCCGCC
101	TGCTCGAAGG	CGGCACCTTA	AATAAGCTCG	CCGCCGACAT	CGGGCTGTTG
151	TCGCAACTGG	GCATCCGACT	CGTCTCATC	CACGCGCGT	ACCACTTCCT
201	CGAaccCCTC	GCCGCCGCGC	AAGgccGCAC	GCCGCATTAT	TGCCGgggtt
251	tGCGCGTTAC	CGACGaAAcc	tcGctcgGAC	AGGCGCAGCA	GtttGCCGGC
301	AccgTCCGCA	GCCGTTTGA	agcCGCATTG	tgcggcagCG	tttcaggatt
351	cgcgCGCGCG	CCTTCCGTCC	CGCTCGTatc	gggcaacttc	ctgacCGCCC
401	GTCcgatggg	cgtgattgac	ggaACCGata	tggataacgc	gggggttatc
451	cgcaaaaaccg	ACACCGCCGC	CCTCCGTTTC	CAACTCGACG	CGGCAATAT
501	CGTCTGGATG	CCGCCGCTCG	GGCATTCTTA	CGGCGGCAAA	ACCTTCAATC
551	TCGATATGGT	GCAGGCGGCC	GCTTCCGTCG	CGTCTCGCT	TCAGGCCGAA
601	AAACTCGTTT	ACCTGACCCT	TTCAGACGGC	ATTCCCCGCC	CCGACGGCAC
651	GCTCGCCGAA	ACCCTCTCGG	CACAGGAAGC	GCAATCGCTG	GCGGAACACG
701	CCGCCAGCGA	AACCCGACGA	CTGATTTCGT	CCGCCGTTGC	CGCGCTCGAA
751	GGCGGCGTGC	ATCGGTCCA	AATCCTCAAC	GGGGCCGCCG	ACGGCAGCCT
801	GCTGCAAGAA	CTCTTCAACC	GCAACGGCAT	CGGCACGTCC	ATTGCCAAAG
851	AAGCCTTCGT	CTCCATCCGG	CAGGCGCACA	GCGGCGACAT	CCCGCACATC
901	CGCGCCCTCA	TCCGCCCGCT	GGAAGAACAG	GGCGTCCTAT	TGCACCGCAG
951	CCGGAATAC	CTCGAAAACC	ACATTTCGGA	ATTTTCATC	CTCGAACACG


```

1001  ACGGCGACCT GTACGGCTGT GCCGCACTCA AAACCTTTGC CGAAGCCGAT
1051  TGC GGCGAAA TCGCCTGCCT TGCCGTCTCG CCGCAGGCAC AGGACGGCGg
1101  ctACGGCGAA CGCCTGCTTG CCCACATTAT CGATAAGGCG CGCGGCATAG
1151  GCATAAGCAG GCTGTTTCGA CTGTCCACAA ATACCGGCGA ATGGTTTGCC
1201  GAACGCGGCT TTCAGACGGC ATCGGAAGAC GAGCTGCCCC AAACGCGGCG
1251  CAAAGACTAC CGCAGCAACG GACGAAACCC GCATATTCTG GTGCGTCGCC
1301  TGCACCGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 148; ORF 040.ng>:

```

g040.pep
1  MNAPDSFVAH FREAAPYIRQ MRGTTLVAGI DGRLLGGTLL NKLAADIGLL
51  SQLGIRLVLI HGAYHFLDRL AAAQGRTPHY CRGLRVTDET SLGQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPMGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWM PPLGHSYGGK TFNLDMVQAA ASVAVSLQAE
201 KLVYLTLSDG ISRPDGTAE TLSAQEAQSL AEHAASETRR LISSAVAALAE
251 GGVHRVQILN GAADGSLLEQ LFTRNGIGTS IAKEAFVSIR QAHSGDIPHI
301 AALIRPLEEQ GVLLHRSREY LENHISEFSI LEHDGDLYGC AALKTFEAD
351 CGEIACLAVS PQAQDGGYGE RLLAHIIDKA RGIGISRLFA LSTNTGEWFA
401 ERGFQTASED ELPETRRKDY RSNGRNPHIL VRRLHR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 149>:

```

m040.seq
1  ATGAGCGCGC CCGACCTCTT TGTCGCCAC TCCCGGAAG CCGTCCCCTA
51  CATCCGCCAA ATGCGCGGCA AAACGCTGGT CGCCGGCATA GACGACCGCC
101 TGCTCGAAGG TGATACCTTA AACAGCTCG CCGCCGACAT CGGGCTGTTG
151 TCGCAACTGG GCATCAGGCT CGTCTCATC CACGGCGCGC GCCACTTCCT
201 CGACCGCCAC GCCGCCGCTC AAGGCCGCAC GCCGCATTAT TGCCGGGGCT
251 TGCGCGTTAC CGACGAAACC TCGCTCGAAC AGGCGCAGCA GTTTGCCGGC
301 AACGTCGCGA GCGGTTTGA AGCCGATTG TGCGGCAGCG TTTCCGGGTT
351 CGCGCGCGCG CCTTCCGTCC CGCTCGTATC GGGCAACTTC CTGACCGCCC
401 GTCCGATAGG TGTGATTGAC GGAACGATA TGGAAATACG GGGCGTTATC
451 CGCAAAACCG ACACCGCCGC CCTCCGTTTC CAACTCGACG CGGGCAATAT
501 CGTCTGGCTG CCGCGCTCG GACATTCCTA CAGCGGCAAG ACCTTCTATC
551 TCGATATGCT TCAAACCGCC GCCTCCGCGC CCGTCTCGCT TCAGGCCGAA
601 AAACGCTGTT ACCTGACCTT TTCAGACGGC ATTTCCCGCC CCGACGGCAC
651 GCTCGCCGAA ACCCTCTCGG CACAGGAAGC GCAATCGCTG GCGGAACACG
701 CCGCGGGGCA AACGCGACGG CTGATTTCGT CCGCCGAAC CTTCACCCGC
751 AACGGCATCG GCACGTCCAT TGCCAAAGAA GCCTTCGTCT CCATCCGGCA
801 rGCGCAywgG G.CGACATCC CGCACATCGC CGCCCTCATC CGCCCGCTGG
851 AAGAACAGGG CATCTGCTG CACCGCAs.c GCGAATACCT CGAAAACCAC
901 ATTTCCGAAT TTTCCATCCT CGAACACGAC GGCAACCTGT ACGGTTGCGC
951 GCCCCTGAAA ACCTTTGCCG AAGCCGATTG CGGCGAAATC GCCTGCCTTG
1001 CCGTCTCGCC GCag.cACAG GACGGCGGCT ACGGCGAACG CnTGCTTGCC
1051 CACATTATCG ATAAGGCGCG CGGCATAGGC ATAAGCAGGC TGTTGCACT
1101 GTCCACAAAT ACCGGCGAAT GGTTTGCCGA ACGCGGCTTT CAGACGGCAT
1151 CGGAAGACGA GTTGCCCGAA ACGCGGCGCA AAGACTACCG CAGCAACGGA
1201 CGGAACCTCG ATATTCTGGT ACGTCGCCTG CACCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 150; ORF 040>:

```

m040.pep
1  MSAPDLFVAH FREAVPYIRQ MRGKTLVAGI DDRLLEGDTL NKLAADIGLL
51  SQLGIRLVLI HGARHFLDRH AAAQGRTPHY CRGLRVTDET SLEQAQQFAG
101 TVRSRFEAAL CGSVSGFARA PSVPLVSGNF LTARPIGVID GTDMEYAGVI
151 RKTDTAALRF QLDAGNIVWL PPLGHSYSGK TFYLDMLQTA ASAAVSLQAE
201 KLVYLTLSDG ISRPDGTAE TLSAQEAQSL AEHAGGQTRR LISSAELFTR
251 NIGITSIAKE AFVSIRQAHX XDIPHIAALI RPLEEQGILL HRXREYLENH
301 ISEFSILEHD GNLYGCAALK TFAEADCGEI ACLAVSPQXQ DGGYGERXLA
351 HIIDKARGIG ISRLFALSTN TGEWFAERGF QTASEDELPE TRRKDYRSNG
401 RNSHILVRRLL HR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 151>:

a040.seq

a040.pcp

1	MIVPDLFVAH	FREAAPYIRQ	MRGKTLVAGI	DDRLLGDTL	NKFAADIGLL
51	<u>S</u> QLGIRLVLI	HGARHFLDRH	AAAQGRTPHY	CRGLRVTDET	SLEQAQQFAG
101	TVRSRFEAAL	CGSVLSGFARA	PSVPLVSGNF	LTPARPIVID	GTDMEVYAGVI
151	RKTDTAALRF	QLDAGNIWVL	PPLGHSYSGK	TFHLDMLQTA	ESAVMSVLQAE
201	KLVLVLTSDG	ISRPDGTLLV	TLGSAEQASL	AEHAGGETTR	LISSAVAALE
251	GGVHRVQILN	GAADGSLQEE	LFTNRNGIGTS	IAKEAFVSIR	QAHSGDIPHI
301	AALIRPLEEQ	GILLHRSREY	LENHISEFSI	LEHDGNLYGC	AALKTFEAD
351	CGEIAPLAVS	PQAQDGGGYE	RLLAHIIDKA	RGIGISRLFA	LSTNTGEWFA
401	ERGFQTAESD	ELPETRKRKY	RSNGRNSHIL	VRLRLR*	

	10	20	30	40	50	60
m040.pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDRRLLLEGDTLNKLAADIGLLSQLGIRLVLI					
a040	MIVPDLFVAHFREAAFYIRQMRGKTLVAGIDRRLLLEGDTLNKFAADIGLLSQLGIRLVLI					
	10	20	30	40	50	60
m040.pep	HGARHFVLDLRHAAAQGRTPHYCRGLRVLTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA					
a040	HGARHFVLDLRHAAAQGRTPHYCRGLRVLTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA					
	70	80	90	100	110	120
m040.pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHYSYSGK					
a040	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHYSYSGK					
	130	140	150	160	170	180
m040.pep	TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR					
a040	TFHLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR					
	190	200	210	220	230	240
m040.pep	TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR					
a040	TFHLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR					

229

	190	200	210	220	230	240
m040.pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI 					
a040	LISSAVAALGGVHRVQILNGAADGSLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI 250 260 270 280 290 300					
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS 					
a040	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS 310 320 330 340 350 360					
m040.pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY 					
a040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY 370 380 390 400 410 420					
m040.pep	RSNGRNSHILVRRLHRX 					
a040	RSNGRNSHILVRRLHRX 430					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 040 shows 88.3% identity over a 436 aa overlap with a predicted ORF (ORF 040.ng) from *N. gonorrhoeae*:

m040/g040

m040.pep	MSAPDLFVAHFREAVPYIRQMRGKTLVAGIDDRLLEGDTLNKLAADIGLLSQLGIRLVLI	60
g040	MNAPDSFVAHFREAAAPYIRQMRGTTLVAGIDGRLLEGGTLNKLAADIGLLSQLGIRLVLI	60
m040.pep	HGARHFLDRHAAAQGRTPHYCRGLRVTDETSLEQAQQFAGTVRSRFEAALCGSVSGFARA	120
g040	HGAYHFLDRLAAAQGRTPHYCRGLRVTDETSLGQAQQFAGTVRSRFEAALCGSVSGFARA	120
m040.pep	PSVPLVSGNFLTARPIGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWLPPLGHSYSGK	180
g040	PSVPLVSGNFLTARPMGVIDGTDMEYAGVIRKTDTAALRFQLDAGNIVWMPPLGHSYGGK	180
m040.pep	TFYLDMLQTAASAAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAGGQTRR	240
g040	TFNLDMVQAAASVAVSLQAEKLVYLTLSDGISRPDGTLAETLSAQEAQSLAEHAASETRR	240
m040.pep	LISSA-----ELFTRNGIGTSIAKEAFVSIRQAHXXDIPHI	276
g040	LISSAVAALGGVHRVQILNGAADGSLQELFTRNGIGTSIAKEAFVSIRQAHSGDIPHI	300
m040.pep	AALIRPLEEQGILLHRXREYLENHISEFSILEHDGNLYGCAALKTFAEADCGEIACLAVS	336
g040	AALIRPLEEQGVLLHRXREYLENHISEFSILEHDGDLYGCAALKTFAEADCGEIACLAVS	360
m040.pep	PQXQDGGYGERXLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY	396
g040	PQAQDGGYGERLLAHIIDKARGIGISRLFALSTNTGEWFAERGFQTASEDELPEPTRRKDY	420
m040.pep	RSNGRNSHILVRRLHRX	413
g040	RSNGRNPHILVRRLHRX	437

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 153>:

```
g041.seq
1   ATGAGTTCGC CCAAACACAT CGGCTTGCG AGCGGCAGCA ACGGCGGCCT
51  GATTACCGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATC GGTGCGCTGG
101 TGTGCGAAGT ACCGCTGACC GATATGATCC GTTATCCGCT GCTGTCCGCC
151 GGTTCAGTTT GGACGGACGA ATACGGCAAT CCGCAGAAAT ACGAAGCCTG
201 CAAACGCCGG CTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCACTC ATTACCACCA GCCTCAGCGA CGACCGCGTC
301 CATCCCGCCC ACGCGCTCAA ATTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCCAACGCGA ATCCGCCGAC AAACCTCGCT GCGTGTGCT GTTTTTGAAA
451 GAATTTTGG GATAA
```

This corresponds to the amino acid sequence <SEQ ID 154; ORF 041.ng>:

```
g041.pep
1   MSSPKHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQYEACKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQRESAD KLACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 155>:

```
m041.seq
1   ATCAGTTCGC CCGAACACAT CGGCTTGCG AGCGGCAGCA ACGGCGGACT
51  GATTACTGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATC GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAGCTT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCCGC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CCCAACGCGA ATCCGCCGAC GAACCTCGCT GCGTCTTGCT GTTTTTGAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 156; ORF 041>:

```
m041.pep
1   ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSQSW LYSPDGGGHT GNGTQRESAD ELACVLLFLK
151 EFLG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 157>:

```
a041.seq
1   ATCAGTTCGC CCGAACACAT CGGCTTGCG AGCGGCAGCA ACGGCGGACT
51  GATTACTGCC GCCGCCTTCG TCGCGGAACC GCAAAGCATA GGCGCGCTGG
101 TGTGCGAAGT GCCGCTGACC GACATGATCC GTTATCCGCT GCTCTCCGCC
151 GGTTCAGCTT GGACAGACGA ATACGGCAAT CCGCAAAAAT ACGAAGTCTG
201 CAAACGCCGG TTGGGCGAAT TGTGCGCGTA TCACAATCTT TCAGACGGCA
251 TCGATTATCC GCCCGCGCTC ATTACCACCA GCCTGTCCGA CGATCGCGTC
301 CATCCCGCCC ACGCGCTCAA GTTCTACGCC AAACCTGCGG AAACCTCGCC
351 GCAATCTTGG CTCTACTCGC CTGACGGCGG CGGCCATACC GGCAACGGCA
401 CGCAGCGCGA AGCCGCCGAC GAACCTCGCT GCGTGTGCT GTTTTTGAAA
451 GAGTTTTTGG GCTAA
```

This corresponds to the amino acid sequence <SEQ ID 158; ORF 041.a>:

```
a041.pep
1   ISSPEHIGLQ GGSNGGLITA AAFVREPQSI GALVCEVPLT DMIRYPLLSA
51  GSSWTDEYGN PQYEVCKRR LGELSPYHNL SDGIDYPPAL ITTSLSDDRV
101 HPAHALKFYA KLRETSPQSW LYSPDGGGHT GNGTQREAAD ELACVLLFLK
```

231

151 EFLG*

m041/a041 98.7% identity over a 154 aa overlap

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
a041	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCRRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSQSW					
a041	PQKYEVCRRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSQSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX					
a041	LYSPDGGGHTGNGTQREAADELACVLLFLKEFLGX					
	130	140	150			

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*ORF 041 shows 96.8% identity over a 154 aa overlap with a predicted ORF (ORF 041.ng) from *N. gonorrhoeae*:

m041/g041

	10	20	30	40	50	60
m041.pep	ISSPEHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	:					
g041	MSSPKHIGLQGGSNGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m041.pep	PQKYEVCRRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSQSW					
g041	PQKYEACRRRLGELSPYHNLSGDIDYPPALITTSLSDDRVPAPAHALKFYAKLRETSPQSW					
	70	80	90	100	110	120
	130	140	150			
m041.pep	LYSPDGGGHTGNGTQRESADELACVLLFLKEFLGX					
g041	LYSPDGGGHTGNGTQRESADKLACVLLFLKEFLGX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 159>:

g041-1.seq

```

1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAAACAA CGACAAGCGC CGCGCACTTT CAGACGGCAT TTTGAATCAA
151 ATGCAGGACA CGCGGCAGAT TCCGTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTC CATCAGAAATG CGGAATATCC GAAGGGCGTG TACCGCATGT
251 GTACGGCGGC GACCTACCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTTCATGA GTTGCTCGGC GACGATGTGT ATTTGGGCGG
351 CGTGTCCGAC TTGGTGGAGC AGCCCAACCG CGCGTGCTG ACTTTGAACA
401 AATCGGGCGG CGATACGGCG TATACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAATTGGTAG AGGGCGGTTT TCACTTCCG GCAGGCAAAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
551 AGTTGACCGA ATCGGGCTAT CCGCGCGAAG TGTGGCTGGT GGAACGCGGC

```

```

601 AAGAGTTTCG AGGAAAGCCT GCCGGCGTAC CAAATCGATA AAGGCGCGAT
651 GATGGTAAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AGACGTATTT GCAGGTGTCTG
751 TCCGAAGGCG GGGCGAAACC GTTGAACTG CTAATGATT GCGATGTGGT
801 CGGCTATCTG GCGGGACATC TTTTGCTGAC GCTGCGCAAG GACTGGCACC
851 GCGCGAACCA AAGCTATCCG AGTGGCGCGT TGGTGGCGGT GAAACTGAAT
901 CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951 GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCAAGCCTGC
1001 TGGAGAATGT ACAAGGCCGT CTGAAAGCGT GGCGGTTTGC CGACAGCAAA
1051 TGGCAGGAAG CCGAGTTGCC GCACCTGCCC TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGCGGCG ACGTGGTTTA TCTTGCCGCC AGCGATTTCA
1151 CCACGCGCGT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAAGTACC
1201 GTCATGCGCC TCCAGCCGCA GCAGTTTGT TCAGACGGCA TCGAAGTGCG
1251 GCAGTTTGG GCGGTGTCGT CCGACGGCGA ACGCATTCCCT TATTTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACACGCCGA CCTTAGTCTA TGCTTACGGA
1351 GGTTCGGCA TTCCTGAATT GCCGATTAT CTGGGCAGCG TCGGCAATA
1401 TTGGCTGGAA GAGGCGAATG CCTTTGTATT GGCAAACATC CGCGGCGGCG
1451 GAGAATTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAC
1501 AAAAGCGTTG ATGATTGTGT GGCAGTCGT CGTGATTTGT CCGAACGCGG
1551 CATGAGTTCG CCCAAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGCC
1601 TATGATTATC CGCCGCTTC GTGCGCGAAC CGCAAAGCAT CGGTGCGCTG
1651 GTGTGCGAAG TACCGCTGAC CGATATGATC CGTTATCCGC TGCTGTCCGC
1701 CGGTTCAAGT TGGACGGACG AATACGGCAA TCCGAGAAA TACGAAGCCT
1751 GCAAACGCCG GCTGGGCGAA TTGTCGCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCGCACT CATTACCACC AGCCTCAGCG ACGACCGCT
1851 CCATCCGCC CACGCGCTCA AATTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CAAACTCGCC TCGGTGTTGC TGTTTTGA
2001 AGAATTTTGG GGATAA

```

This corresponds to the amino acid sequence <SEQ ID 160; ORF 041-1.ng>:

g041-1.pep

```

1 MKSYDPYR FENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILNQ
51 MQDTRQIPFC QEHRARMYHF HQNAEYPKGV YRMCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLNKSNGGDTA YTLEVDLEAG
151 ELVEGGFHF AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPAY QIDKGAMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 SEGGAAPLNL PNDQDVVGYL AGHLLTLRK DWHRANQSY SPALVAVKLN
301 RGEELGAAQL FAPDETQALE SVETTKRFV ASLLENVQGR LKAWRFADSK
351 WQEAELPHLP SGALEMTDQP WGGDVVYLAA SDFTTPTLTF ALDLNVMELT
401 VMRLQPQFV SDGIEVRQFW AVSSDGERIP YFHVGNKNAAP DPTLVYAYG
451 GFGEPELPHY LGSVGKYWLE EGNFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLAVV RDLSEGRMS PKHIGLQGS NGGLITAAAF VREPQSIGAL
551 VCEVPLDMI RYPLLSAGSS WTDEYGNPQK YEACKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDRVHFA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQRESADKLA CVLLFLKEFL G*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 161>:

m041-1.seq

```

1 ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAACCG TCGATTCCGC
51 CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAGAAAA CGACAAGGCG CGCGCGCTTT CAGACGCGAT TTTGGCGCAG
151 TTGCAGGACA CGCGGCAGAT TCCGTTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAGGACG CGGAGTATCC GAAGGGCGTG TACCGCGTGT
251 GTACCGCGGC GACGTATCGT TCCGGCTATC CCGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTGACGCA ATGCTTGGC GACGATGTGT ATTTGGCGCG
351 CGTGTGCGAC TTGGTGGAAC AGCCCAACCG CGCGTTGTTA AACTGAGCA
401 AATTGGGCGC CGATACGGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAGTTGGTCG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 TGGCGCGCAT GAAACAGCG TGTGGGTGTG TCCGGCTTGG AACGAACGCC
551 AGTTGACCCA ATCGGGCTAT CCGCGCGAAG TATGGCTGGT GGAACGCGGC
601 AAGAGTTTCG AGGAAAGCCT GCCTGTGTAT CAAATCGGCG AAGACGGCAT
651 GATGGTGAAC GCGTGGCGTT ATCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCGGACGGT TTTTACACCA AAACCTATTT GCGGGTCTCA

```

```

751  GCCGAAGGCG AGGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTGGT
801  CGGCTATCTG GCGGGGCATC TTTTGCTGAC GCTGCGCAAG GACTGGAACC
851  GCGCGAACCA AAGCTATCCG AGCGGCGCGC TGGTGGCGGT GAAGCTGAAT
901  CGGGGCGAAC TCGGGGCGGC GCAGCTTTTG TTTGCGCCCG ATGAAACGCA
951  GGCATTGGAA AGCGTGGAAA CGACCAAGCG TTTTGTGGTG GCGAGCCTGT
1001 TGGAGAACGT ACAAGGCCGT CTGAAAGCAT GCGGTTTTCG CGACGGCAAA
1051 TGGCAGGAAG TCGAATTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCT TGGGGCGGCG ACGTGGTTTA CCTTGCCGCC AGCGATTTC
1151 CCACGCGCT GACGCTGTTT GCGCTGGATT TGAACGTGAT GGAAGTACG
1201 GTCATGCGCC GCCAGCGCA GCAGTTGAT TCAGACGGCA TTAACGTGCA
1251 GCAGTTTTCG ACGACTTCGG CTGACGGCGA GCGCATTCTT TATTCCACG
1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCTACGGC
1351 GGTTCGCGCA TTCCCGAATT GCCGCATTAT CTGGGCAGCA TTGGCAATA
1401 TTGGCTGGAA GAGGCAATG CCTTTGTATT GCGCAACATC CGCGGCGGCG
1451 GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTTATT GGCAGTCGTG CCGGATTTGT CCGAACGCGG
1551 TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601 TGATTACTGC CGCCGCTTC GTGCGCGAAC CGCAAAGCAT CGGCGCGCTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAAACGCC GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCGCGCTC CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCCG
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACCCAACGCG AATCCGCCGA CGAACTCGCC TCGCTCTTGC TGTTTTGA
2001 AGAGTTTTTG GGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 162; ORF 041-1>:

m041-1.pep

```

1  MKSYDPYRHFENLDSAETQ NFAAEANAET RARFLENDKA RALSDGILAQ
51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKLGS DTA YLEVDLEAG
151 ELVEGGFHF AGKNHVS WRD ENSVWVCPAW NERQLTQSGY PREVWLVERG
201 KSFEESLPVY QIGEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLRVS
251 AEGEAKPLNL PNDGDVVGYL AGHLLLT LRK DWNRRANQSY SGALVAVKLN
301 RGELGAAQLL FAPDETQALE SVETTKRFV ASLLENVQGR LKAWRFADGK
351 WQEVLEPLRLP SGALEMTDQP WGGDVVYLAA SDFTTPLTLF ALDLNVMELT
401 VMRRQPQQFD SDGINVQQFW TTSADGERIP YFHVGNNAAP DMPTLVYAYG
451 GFGIPELPHY LGSIGKYWLE EGNAAFVLANI RGGGEFGPRW HQAAGQISKH
501 KSVDDLAVV RDLSEGRIS PEHIGLQGS NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLS DG
601 IDYPPALITT SLSDDRVHFA HALKFYAKLR ETSQSWLYS PDGGGHTGNG
651 TQRESADELA CVLLFLKEFL G*

```

m041-1/g041-1 94.6% identity in 671 aa overlap

	10	20	30	40	50	60
m041-1.pep	MKSYDPYRHFENLDSAETQ	NFAAEANAET	RARFLENDKA	RALSDGILAQ	LQDTRQIPFC	
g041-1	MKSYDPYRHFENLDSAETQ	NFAAEANAET	RARFLNNDKA	RALSDGILNQ	MDTRQIPFC	
	10	20	30	40	50	60
m041-1.pep	QEHRARMYHFHQDAEY	PKGVYRVCTAATYR	SGYPEWKILF	SVADFDELLG	DDVYLGGVSH	
g041-1	QEHRARMYHFHQDAEY	PKGVYRMCTAATYR	SGYPEWKILF	SVADFDELLG	DDVYLGGVSH	
	70	80	90	100	110	120
m041-1.pep	LVEQPNRALLTLSKLGS	DYATLEVLEAGEL	VEGGFHF	PAGKNHVS	WRDENS	SVWVCPAW
g041-1	LVEQPNRALLTLNKS	GGDTAYTLEVLEAGEL	VEGGFHF	PAGKNHVS	WRDENS	SVWVCPAW
	130	140	150	160	170	180
m041-1.pep	NERQLTQSGYPRE	VWLVERGK	SFEESLPVY	QIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG
g041-1	NERQLTQSGYPRE	VWLVERGK	SFEESLPVY	QIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG
	190	200	210	220	230	240
m041-1.pep	NERQLTQSGYPRE	VWLVERGK	SFEESLPVY	QIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG
g041-1	NERQLTQSGYPRE	VWLVERGK	SFEESLPVY	QIGEDGMMVN	AWRYLDPQGS	PIDLIEASDG

g041-1	DERQLTESGYPREVWLVERGKSFEESLPAYQIDKGAMMVNAWRYLDPQGSPIDLIEASDG
	190 200 210 220 230 240
m041-1.pep	250 260 270 280 290 300
	FYTKTYLRVSAEGEAKPLNLPNDCDVVGYLGHLLTLRKDWNRRANQSYPSGALVAVKLN
	:
g041-1	FYTKTYLQVSSEGGAKPLNLPNDCDVVGYLGHLLTLRKDWHRRANQSYPSGALVAVKLN
	250 260 270 280 290 300
m041-1.pep	310 320 330 340 350 360
	RGELGAAQLLFAPDETOALESVETTKRFVVASLLENVQGRLLKAWRFADGKWQEVLPRLP
	:
g041-1	RGELGAAQLLFAPDETOALESVETTKRFVVASLLENVQGRLLKAWRFADSKWQEAELPHLP
	310 320 330 340 350 360
m041-1.pep	370 380 390 400 410 420
	SGALEMTDQPWGGDVVYLAASDFTTPTLTFALDLNVMELTVMRRQPQQFSDGINVQQFW
	:
g041-1	SGALEMTDQPWGGDVVYLAASDFTTPTLTFALDLNVMELTVMRLQPQQFVSDGIEVRQFW
	370 380 390 400 410 420
m041-1.pep	430 440 450 460 470 480
	TTSADGERIPYFHVGNKNAAPDMPTLVYAYGGFGIPELPHYLGSGIKYWLEEGNAFVLANI
	:::
g041-1	AVSSDGERIPYFHVGNKNAAPDPTLVYAYGGFGIPELPHYLGSGIKYWLEEGNAFVLANI
	430 440 450 460 470 480
m041-1.pep	490 500 510 520 530 540
	RGGGEFGPRWHQAAQGISKHKSVDDLAVVRDLSESGISSPEHIGLQGGSNGLITAAAF
	:
g041-1	RGGGEFGPRWHQAAQGISKHKSVDDLAVVRDLSESGMSSPKHIGLQGGSNGLITAAAF
	490 500 510 520 530 540
m041-1.pep	550 560 570 580 590 600
	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCRRRLGELSPYHNLSDG
	:
g041-1	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEACKRRRLGELSPYHNLSDG
	550 560 570 580 590 600
m041-1.pep	610 620 630 640 650 660
	IDYPPALITTSLSDDRVPALHAKFYAKLRETSQSWLYSPDGGGHTGNGTQRESADELA
	:
g041-1	IDYPPALITTSLSDDRVPALHAKFYAKLRETSQSWLYSPDGGGHTGNGTQRESADKLA
	610 620 630 640 650 660
m041-1.pep	670
	CVLLFLKEFLGX
g041-1	CVLLFLKEFLGX
	670

m041-1/P55577

sp|P55577|Y4NA_RHISN PROBABLE PEPTIDASE Y4NA >gi|2182536 (AE000086) Y4na [Rhizobium sp. NGR234] Length = 726

Score = 370 bits (940), Expect = e-101

Identities = 217/682 (31%), Positives = 331/682 (47%), Gaps = 22/682 (3%)

Query: 2 KSYDPYRHFENLDSAEQNFQAEANAETRARFLENDKARALSDGILAQLDTRQIPFCQ 61
K DP + +D + + N T + ++ + L LQ T +I

Sbjct: 42 KDASDPRAYLNEIDGDKAMTWVEAHNLSTVDKLSKDFRYSEYQADALTILQATDRIASPS 101

Query: 62 EHRARMY-HFHQDAEYPKGVYRVCTAATYRSGYPEWKILFSVADFDELLGDDVYLGGVSH 120
R M +F QD + +G++R T +YRSG P+W+ + V + G G

Sbjct: 102 FARDGMIDNFWQDGHVQGLWRRRTTWESYRSGNPQWRTILDVDALSKEGKTWVFEGGDC 161

Query: 121 LVEQPNRALLTSLKSGSDTAYTLEVDLEAGELVEGGFHFAGKNHVSWRDENSVMVCPAW 180
L N L+ LS G D E D+ GE V+ GF P GK V+W DEN++V W

Sbjct: 162 LPPTSNLCLIRLSDGGKDADVREFDIAKGEFVKEGFVLPEGKQSVTWVDENTIIYVTREW 221

Query: 181 NERQLTQSGYPREVWLVERGKSFEESLPVYQ-----IGEDGMM--VNAWRYLDPQGSPI 232
 ++T SGY ++V+RG+S ++++ +++ E G++ ++ +D +

Sbjct: 222 TPGEVTSSSGYAYVTKVVKRQSLDQAVEIFRGQKQKDVSAERGVLRLDIDGKYVMDTSYRGL 281

Query: 233 DLIEASDGFYTKTYLRVSAEGEAKPLNLPNDCDVVGYLAGHLLTLRKDOWNRANQS-YPS 291
 D FY + + L LP GY G + L+ DW A + + +

Sbjct: 282 DFFNTELAIFYPNH----PDTRKVVLPPLTTAVFSGYKQAIYWLKSDWTSAGKTVFHN 337

Query: 292 GALVAVKLNRLGELGAAQL----LFAPDETQALESVETTKRFVVASLLENVQGRKAWRFA 347
 GA++A L A++ LF P+E Q++ TK +V S+L NV +++++ F

Sbjct: 338 GAIIAFDLKAALADPARVEPLVLFMPNEHQSVAGTTQTKNRLVLSILSNVTSEVRSFDFG 397

Query: 348 DGKWQEVLPRLPSGALEMTDQFWGGDVVYLAASDETTPLTLFALDLNVMELTVMRRQPQ 407
 G W +L + L +T D +++ + F P TLF D ++ + P

Sbjct: 398 KGWSSFKLALPENSTLSLTSSDDESQLFVVFSEGFLEPSTLFCADAATGQVEKITSTPA 457

Query: 408 QFDSGGINVQQFWTTSADGERIPYFHVGNKNAAP---DMPTLVYAYGGFGIPELPHYLGSI 464
 +FD+ G+ QQFW TS DG ++PYF V + PT++YAYGGF IP P Y +

Sbjct: 458 RFDAGGLQAQQFWATSKDGTKVPYFYLVARQDVKLDGNTPTILYAYGGFGIPELPHYLGSI 517

Query: 465 GKYWLEEGNAFLANIRGGGEFGPRWHQAAQGISKHKSVDDLLAVVRDLSERGISSPEHI 524
 GK WLE+G A+ LANIRGGGEFGP+WH A ++ + DD AV +DL + ++S H+

Sbjct: 518 GKWLWLEKGAALANIRGGGEFGPKWHDAGLKTNRQVYDDFQAVQDLIAKVTSTPHL 577

Query: 525 GLQGGSSNGGLITAAAFVREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPKQYEV 584
 G+ GGSNGGL+ ++ P A+V +VPL DM+ + +SAG+SW EYG+P V

Sbjct: 578 GIMGGSNGGLLMGVQMIQRPDLWNAVVIQVPLDMVNFTRMSAGASWQAEYGPDD-PVE 636

Query: 585 KRRLGELSPYHNLSGDIDYPPALITTSLSDDRVHPAHALKFYAKLRETSAQSWLYSPDGG 644
 L +SPYHN+ G+ YP TS DDRV P HA K A + + Y G

Sbjct: 637 GAFLRSISPYHNKAGVAYPEPFETSTKDDRVGPVHARKMAALFEDMGLPFYYENIEG 696

Query: 645 GHTGNGTQRESADELACVLLFL 666
 GH +E A A +++

Sbjct: 697 GHAAAANLQEHARRYALEYIYM 718

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 163>:

a041-1.seq

```

1  ATGAAATCCT ACCCCGACCC CTACCGCCAT TTTGAAAACC TCGATTCCGC
51  CGAAACGCAA AACTTCGCTG CTGAAGCGAA TGCCGAAACG CGCGCGCGTT
101 TTTTAAACAA CGACAAGGCA CGCGCATTGT CTGACGGCAT TTTGGCGCAG
151 TTGCAGGACA CGCGGCAAAT TCCGTTTGT CAGGAACACC GCGCGCGGAT
201 GTACCATTTT CATCAAGATG CGGAATATCC GAAAGGCGTG TACCGCGTGT
251 GTACCGCGGC GACTTACCGT TCGGGCTATC CTGAGTGGAA AATCCTGTTT
301 TCGGTGGCGG ATTTGACGCA ATTGCTCGGT GACGATGTAT ATCTAGGCGG
351 CGTGTGCGAC CTGGTGGAA CAGCCCAACCG CGCGTGTGTA AACTGAGCA
401 AATCGGGCGG CGATACCGCG TACACGCTGG AAGTGGATTT GGAAGCAGGG
451 GAGTTGGTAG AAGGCGGTTT TCACTTTCCG GCAGGCAAAA ACCATGTGTC
501 GTGGCGCGAT GAAAACAGCG TGTGGGTGTG TCCGGCTTGG GACGAACGCC
551 AGTTGACCGA ATCGGGCTAT CCGCGCGAGG TGTGGCTGGT GGAACGCGGC
601 AAGAGTTTCG AGGAAAGCCT GCCGGTGAT CAAATTGCTG AAGACGGCAT
651 GATGGTGAAC GCGTGGCGTT ACCTCGATCC GCAGGGTTCG CCGATTGATT
701 TGATTGAAGC GTCTGACGGT TTTTACACCA AAACCTATTT GCAGGTCTCA
751 GCCGAAGGCG AAGCGAAACC GTTAAACCTG CCCAACGATT GCGACGTAGT
801 CGGCTATCTG GCCGGACATC TTTTGCTGAC CTTGCGTAAA GACTGGCACC
851 GCGCGAACCA AAGCTATCCG AGTGGCGCAT TGGTAGCAGT AAAATTAAC
901 CGCGGCGAAT TGGGCGCGGC GCAGCTTTTG TTTGCGCCCA ATGAAACGCA
951 GGCATTGGAA AGCGTGGA ACGACCAAGCG TTTTGTCTGT GCGAGCCTGC
1001 TGGAAAACGT ACAGGTCGT CTGAAAGCGT GGCGTTTTAC TGATGGCAAA
1051 TGGCAGGAAA CCGAGTTGCC GCGCCTGCCT TCGGGCGCGT TGGAAATGAC
1101 CGACCAACCG TGGGGGGGCG ACGTAGTTTA CCTTGCCTCG AGCGATTTC
1151 CCACGCCGCT GACGCTGTTT GCATTGGATT TGAACGTGAT GGAACGTACC
1201 GTCATGCGCC GCCAGCGCA GCAGTTGAT TCAGACGGCA TTAACGTGCA
1251 GCAGTTTGG ACGACTTCGG CTGACGGCGA GCGCATTCCT TATTTCACG
1301 TCGGCAAAAA CGCCGCGCCC GACATGCCGA CGCTGGTCTA TGCCTACGGC
1351 GGTTTCGGCA TTCCCGAATT GCCGATTAT CTGGGCAGCA TTGGCAATA
1401 TTGGCTGGAA GAGGGCAATG CCTTGTATT GCGAACATC CGCGGCGGCG

```

```

1451 GCGAGTTCGG CCCGCGCTGG CATCAGGCGG CGCAGGGAAT CAGCAAACAT
1501 AAAAGCGTTG ATGATTATT GGCAGTCGTG AGCGATTTGT CCGAACGCGG
1551 TATCAGTTCG CCCGAACACA TCGGCTTGCA GGGCGGCAGC AACGGCGGAC
1601 TGATTACTGC CGCCGCCTTC GTGCGCGAAC CGCAAAGCAT AGGCGCGCTG
1651 GTGTGCGAAG TGCCGCTGAC CGACATGATC CGTTATCCGC TGCTCTCCGC
1701 CGGTTCAAGC TGGACAGACG AATACGGCAA TCCGCAAAAA TACGAAGTCT
1751 GCAAACGCGG GTTGGGCGAA TTGTCGCCGT ATCACAATCT TTCAGACGGC
1801 ATCGATTATC CGCCCGCGCT CATTACCACC AGCCTGTCCG ACGATCGCGT
1851 CCATCCCGCC CACGCGCTCA AGTTCTACGC CAAACTGCGC GAAACCTCGC
1901 CGCAATCTTG GCTCTACTCG CCTGACGGCG GCGGCCATAC CGGCAACGGC
1951 ACGCAGCGCG AAGCCGCCGA CGAACTCGCC TGCCTGTTGC TGTTTTGTAA
2001 AGAGTTTTTG GGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 164; ORF 041-1.a>:

a041-1.pep

```

1  MKSYDPYRHFENLDSAETQ NFAAEANAET RARFLNNDKA RALSDGILAQ
51  LQDTRQIPFC QEHRARMYHF HQDAEYPKGV YRVCTAATYR SGYPEWKILF
101 SVADFDELLG DDVYLGGVSH LVEQPNRALL TLSKSGGDTA YTLEVDLEAG
151 ELVEGGFHFH AGKNHVSWRD ENSVWVCPAW DERQLTESGY PREVWLVERG
201 KSFEESLPVY QIAEDGMMVN AWRYLDPQGS PIDLIEASDG FYTKTYLQVS
251 AEGEAKPLNL PNDCDVVGYL AGHLLTLRKH DWHRANQSYR SGALVAVKLN
301 RGELGAAQLL FAPNETQALE SVETTKRFVV ASLLENVQGR LKAWRFTDGR
351 WQETELPRLP SGALEMTDQP WGGDVVYLAA SDFTTPTLTF ALDLNVMELT
401 VMRRQPQQFD SDGINVQQWF TTSADGERIP YFHVGNNAAP DMPTLVYAYG
451 VFGIPELPHY LGSIGKYWLE EGNAFVLANI RGGGEFGPRW HQAAQGISKH
501 KSVDDLLAVV SDSLSEGISS PEHIGLQGGG NGGLITAAAF VREPQSIGAL
551 VCEVPLTDMI RYPLLSAGSS WTDEYGNPQK YEVCKRRLGE LSPYHNLSDG
601 IDYPPALITT SLSDRVHHPA HALKFYAKLR ETSPQSWLYS PDGGGHTGNG
651 TQREAADELA CVLLFLKEFL G*

```

a041-1/m041-1 97.9% identity in 671 aa overlap

	10	20	30	40	50	60
a041-1.pep	MKSYDPYRHFENLDSAETQ	NFAAEANAET	RARFLNNDKA	RALSDGILAQ	LQDTRQIPFC	
m041-1	MKSYDPYRHFENLDSAETQ	NFAAEANAET	RARFLNNDKA	RALSDGILAQ	LQDTRQIPFC	
	10	20	30	40	50	60
a041-1.pep	QEHRARMYHFHQDAEYPKGV	YRVCTAATYR	SGYPEWKILF	SVADFDELLG	DDVYLGGVSH	
m041-1	QEHRARMYHFHQDAEYPKGV	YRVCTAATYR	SGYPEWKILF	SVADFDELLG	DDVYLGGVSH	
	70	80	90	100	110	120
a041-1.pep	LVEQPNRALLTLSKSGGDTA	YTLEVDLEAG	ELVEGGFHFH	PAGKNHVSWRD	ENSVWVCPAW	
m041-1	LVEQPNRALLTLSKSGGDTA	YTLEVDLEAG	ELVEGGFHFH	PAGKNHVSWRD	ENSVWVCPAW	
	130	140	150	160	170	180
a041-1.pep	DERQLTESGYPREVWLVERG	KSFEESLPVY	QIAEDGMMVN	AWRYLDPQGS	PIDLIEASDG	
m041-1	DERQLTESGYPREVWLVERG	KSFEESLPVY	QIAEDGMMVN	AWRYLDPQGS	PIDLIEASDG	
	190	200	210	220	230	240
a041-1.pep	FYTKTYLQVSAE	GEAKPLNPNDCDVVGYL	AGHLLTLRKH	DWHRANQSYR	SGALVAVKLN	
m041-1	FYTKTYLQVSAE	GEAKPLNPNDCDVVGYL	AGHLLTLRKH	DWHRANQSYR	SGALVAVKLN	
	250	260	270	280	290	300
a041-1.pep	RGELGAAQLLFAPNETQALE	SVETTKRFVV	ASLLENVQGR	LKAWRFTDGR	WQETELPRLP	
m041-1	RGELGAAQLLFAPNETQALE	SVETTKRFVV	ASLLENVQGR	LKAWRFTDGR	WQETELPRLP	
	310	320	330	340	350	360
a041-1.pep	RGELGAAQLLFAPNETQALE	SVETTKRFVV	ASLLENVQGR	LKAWRFTDGR	WQETELPRLP	
m041-1	RGELGAAQLLFAPNETQALE	SVETTKRFVV	ASLLENVQGR	LKAWRFTDGR	WQETELPRLP	

237

	370	380	390	400	410	420
a041-1.pep	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMVRRQPQQFSDSGINVQQFW					
m041-1	SGALEMTDQPWGGDVVYLAASDFTTPLTLFALDLNVMELTMVRRQPQQFSDSGINVQQFW					
	370	380	390	400	410	420
	430	440	450	460	470	480
a041-1.pep	TTSADGERIPYFHVGNAAAPDMPTLVYAYGGFGIPELPHYLSIGKYWLEEGNAFVLANI					
m041-1	TTSADGERIPYFHVGNAAAPDMPTLVYAYGGFGIPELPHYLSIGKYWLEEGNAFVLANI					
	430	440	450	460	470	480
	490	500	510	520	530	540
a041-1.pep	RGGGEFGPRWHQAAQGISKHKSVDDLAVVSDLSERGISSPEHIGLQGSNGGLITAAAF					
m041-1	RGGGEFGPRWHQAAQGISKHKSVDDLAVVSDLSERGISSPEHIGLQGSNGGLITAAAF					
	490	500	510	520	530	540
	550	560	570	580	590	600
a041-1.pep	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
m041-1	VREPQSIGALVCEVPLTDMIRYPLLSAGSSWTDEYGNPQKYEVCKRRLGELSPYHNLSDG					
	550	560	570	580	590	600
	610	620	630	640	650	660
a041-1.pep	IDYPPALITTSLSDDRVPALHAKFYAKLRETSPQSWLYSPDGGGHTGNGTQREAADELA					
m041-1	IDYPPALITTSLSDDRVPALHAKFYAKLRETSAQSWLYSPDGGGHTGNGTQREAADELA					
	610	620	630	640	650	660
	670					
a041-1.pep	CVLLFLKEFLGX					
m041-1	CVLLFLKEFLGX					
	670					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 165>:

g042.seq

```

1  ATGACGATGA TTTGCTTGCG CTTCCAagcG TTCGTGCCGC ATACCAGCGC
51  GTTATCCAAC ACTTCCACGG CAGCCGGCCC TTCCTGCCCC ATGGCGGCGG
101 TGCGGTCGAT GATGAAAATC CAGCCGGGGT TTTTCTCTTT GATGTATTCC
151 AAGGAAACGG GCTGCCCCGTG CCCTTCGTTG CGTAAAGATT CGTCCACGGG
201 CGGCAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GATTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCTGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG cggCTTCGCG CTTTGGGGCG AACAGCGCGT CAATCTGCGC
351 ATTCAATTCC GCCACGCGCG CTTCTTACC GAAAATCCGC GACAGGGTCT
401 CCATCTGCTT CTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAAA
451 TCTATGgtgG TCGCGTTTTT CGCCAACGTG TCATACGCTT CCGCACCCGG
501 CCCGCCGTA ATGACAAACT GCGGATTGTG GCGGTGCAGG GATTGCAAT
551 CGGGCTCAAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAA TGCAGGCTGC
601 AAATAG

```

This corresponds to the amino acid sequence <SEQ ID 166; ORF 042.ng>:

g042.pep

```

1  MTMICLRFOA FVPHTSALSN TSTAAGPSCP MAAVRSMKI QPGFFSLMYS
51  KETGCPCLSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDSSTPRP
101 LPLAASRFWA NSASICAFNS ATRASLPKIR DRVSICFSPV VRILEPLSTVK
151 SMVVAFFANC SYASAPGPPV MTNCLWRCR DSQSGSNSVP TVAALSNAAG
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 167>:

m042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAmT ACTTCGACAG CCGcCGGCCy TTCyTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCAG GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 168; ORF 042>:

m042.pep

```

1 MTMICLRFQA FVPRTSALSX TSTAAGXSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTSXGLXRCR ASXSGSNSVP TVAALSNAGC
201 K*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 169>:

a042.seq

```

1 ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCAG GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGGTA ATGACAAGCT GAGGATTGTA GCGGTGCAGG GCTTCGTAAT
551 CGGGCTCGAA CAGCGTCCCC ACCGTTGCCG CCTTGTCAA TGCAGGCTGC
601 AAATAA

```

This corresponds to the amino acid sequence <SEQ ID 170; ORF 042.a>:

a042.pep

```

1 MTMICLRFQA FVPRTSALSX TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVTDTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*GL*RCR AS*SGSNSVP TVAALSNAGC
201 K*

```

m042/a042 99.0% identity over a 201 aa overlap

	10	20	30	40	50	60
m042.pep	MTMICLRFQAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
a042	MTMICLRFQAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
	10	20	30	40	50	60
m042.pep	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDTSPRPLAASRVWANSASICAFNS					
a042	RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVTDTSPRPLAASRVWANSASICAFNS					
	70	80	90	100	110	120
m042.pep	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
a042	AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR					
	130	140	150	160	170	180

```

a042      |||||
          AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR
          130      140      150      160      170      180

          190      200
m042.pep  ASXSGSNSVPTVAALSNAGCKX
          |||||
a042      ASXSGSNSVPTVAALSNAGCKX
          190      200

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 042 shows 93.0% identity over a 201 aa overlap with a predicted ORF (ORF 042.ng) from *N. gonorrhoeae*:

m042/g042

```

m042.pep      10      20      30      40      50      60
              MTMICLRFQAFVPRTSALSXTSTAAGXSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL
              |||||:|||||
g042          MTMICLRFQAFVPHTSALSNSTSTAAGPSCPMAAVRSMMKIQPGFFSLMYSKETGCPCPSL
              10      20      30      40      50      60

m042.pep      70      80      90      100     110     120
              RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRVWANSASICAFNS
              |||||:|||||
g042          RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVDSTSPRPLPLAASRFWANSASICAFNS
              70      80      90      100     110     120

m042.pep      130     140     150     160     170     180
              AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANC SYASAPGPPVMTSXGLXRCR
              |:||||| :|||||:|||||:|||||: || |||
g042          ATRASLPKIRDVSICFSPLVRILPLSTVKS MVVAFFANC SYASAPGPPVMTNCGLWRCR
              130     140     150     160     170     180

m042.pep      190     200
              ASXSGSNSVPTVAALSNAGCKX
              | |||||
g042          DSQSGSNSVPTVAALSNAGCKX
              190     200

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 171>:

m042-1.seq

```

1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCGC GTACCAGCGC
51 GTTATCCAAT ACTTCGACAG CCGCCGGCCC TTCCTGCCCG ATGGCGGCGG
101 TACGGTCGAT GATGAAAATC CAATCGGGGT TTTTCTCTTT GATGTATTCTG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCCGT GTATCCAAC TGCCAACCGC GACTGCGTGC
251 CGAAGGCGGA CACCTTGTTG CCCGTAACCG ACAGCACCAG CCCGCGTCCT
301 TTGCCTTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TTCGCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGGTGG TCGCGTTTTT CGCTAACTGT TCATACGCTT CCGCGCCCGG
501 CCCGCCGTA A

```

This corresponds to the amino acid sequence <SEQ ID 172; ORF 042-1>:

m042-1.pep

```

1  MTMICLRFQA FVPRTSALSN TSTAAGPSCP MAAVRSMMKI QSGFFSLMYS
51 KETGCPCPSL RKDSSTGGRP MSPCIQLANR DCVPKADTLL PVDSTSPRP
101 LPLAASRVWA NSASICAFNS AARASLPKIR AKVSICFSPL VRILPLSTVR
151 SMVVAFFANC SYASAPGPPV MTS*

```

m042-1/g042 95.4% identity in 173 aa overlap

		10	20	30	40	50	60
m042-1.pep		MTMICLRFOAFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
g042		MTMICLRFOAFVPHTSALSNTSTAAGPSCPMAAVRSMMKIQPGFFSLMYSKETGCPCPSL					
		10	20	30	40	50	60
		70	80	90	100	110	120
m042-1.pep		RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRVWANSASICAFNS					
g042		RKDSSTGGRPMSPCIQLANRDCVPKADTLLPVT DSTSPRPLPLAASRFWANSASICAFNS					
		70	80	90	100	110	120
		130	140	150	160	170	
m042-1.pep		AARASLPKIRAKVSICFSPLVRILPLSTVRSMVVAFFANCYSYASAPGPPVMTSX					
		:	:				
g042		ATRASLPKIRDRVSIKFSPLVRILPLSTVKSVMVVAFFANCYSYASAPGPPVMTNCGLWRCR					
		130	140	150	160	170	180
g042		DSQSGSNSVPTVAALSNAGCKX					
		190	200				

```
a042-1.seq
1  ATGACGATGA TTTGCTTGCG CTTCCAAGCG TTCGTGCCCG GTACCAGCGC
51  GTTATCCAAT ACTTCGACAG CCGCGGGCCG TTCCTGCCCG ATGGCGGCGC
101 TACGGTCGAT GATCGAAATC CAATCGGGCT TTTTCTCTTT GATGATATCG
151 AAGGAAACAG GCTGCCCGTG CCCCTCGTTG CGTAAAGATT CGTCTACAGG
201 CGGTAGGCCG ATGTCGCGCT GTATCCAACT TGCCAACCGC GACTCGCTGC
251 CGAAGGCGGA CACCTTGTTG CCGTAAACCG ACAGCACCAG CCCGCGTCTC
301 TGGCCTTGG CGGCTTCGCG CGTTTGGGCG AACAGCGCGT CAATCTGCGC
351 CTTCAATTCC GCCGCGCGCG CTTCTTGCC GAAATCCGC GCCAAGGTCT
401 CCATCTGCTT TCGCGCGCTG GTGCGGATAT TGCCGTTGTC CACCGTCAGA
451 TCTATGTTGG TCGGTTTTT CGCCAAGTGT TCATACGCTT CCGGCCCCGG
501 CCGCCGGTA A
```

a042-1.pep

1	MTMICLRFQA	FVPRTSALS	N TSTAAGP	SCP MAAVR	SMMKI QSGFF	SLMYS
51	KETGCP	CPSL RKD	SSTGGRP	MSPCIQL	ANR DCPVK	ADTLL
101	LPLAAS	RVWA NS	SICAFNS	AARASLP	KIR AKVSI	CIFSL
151	SMVVA	FFANC	SYASAP	GGPV	MTS*	

		10	20	30	40	50	60
m042-1.pep		MTMICLRFAQFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
a042-1		MTMICLRFAQFVPRTSALSNTSTAAGPSCPMAAVRSMMKIQSGFFSLMYSKETGCPCPSL					
		10	20	30	40	50	60
		70	80	90	100	110	120
m042-1.pep		RKDSSTGGRMSPCQIQLANRDCVPKADTLLPVTDTSTSPRPLPLAASRVWANSASICAENS					
a042-1		RKDSSTGGRMSPCQIQLANRDCVPKADTLLPVTDTSTSPRPLPLAASRVWANSASICAENS					
		70	80	90	100	110	120
		130	140	150	160	170	
m042-1.pep		AARASLPKIRAKVSICFSPLVRILPLSTVRSMMVVAFFANCYSASAPGPPVMTSX					
a042-1		AARASLPKIRAKVSICFSPLVRILPLSTVRSMMVVAFFANCYSASAPGPPVMTSX					
		130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 175>:

241

```

g043.seq
1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GTCGGCCCAT CAGCACTTTT
51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GCCTGAACGC TTCgttgaAC
101 CGTCCCGCGT ggcggtagcc gcAAAAGTGC ATcGCGGCTT GGATGGTGCT
151 GCCCGATTCTG ATGAGGGcga gcGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 GTCCGGCGAC GGTTCGCGCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTCTGTTCA GCCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGGCG
301 GCGGAATTCG CTGTTCAAAA TATCGGCGGC TTCGTCTATG CGCCGGCGGC
351 GGTAGCCGTT GTCGTGGCGG CGGAAGGTGA AGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 176; ORF 043.ng>:

```

g043.pep
1  MVVSNQNIYA VGPSALFHIR RQKSVMPER FVEPSRVAVA AKVHRGLDGA
51  ARFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFDGDQRT
101 GEFVQNIGG FVYAPAAVAV VVAAEGEA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 177>:

```

m043.seq
1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCT CAGCACTTCT
51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101 CGTCCCGCGT GCGGCTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT
151 GCCCGATTCTG ATGAGGGCGA GCGCGTGTTT CAGCCGCAGG CGGCGCAGGC
201 ATCCGGCGAC GGTTCGCGCG GTTTGCGCTT TGAAATAGCG TTTCAGGTAG
251 CATTCTGTTCA GTCCGACGCG GCGGGCGATT TCGGCGATGG TCAGCGGACG
301 GCGGAATTCG TGTTGAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGCG
351 GGTAAACGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 178; ORF 043>:

```

m043.pep
1  MVVSNQNIYA AGPSALLHIR RQKSVMPER FVEPSRVAVA AKVHGGLDGA
51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFDGDQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* *menA* with *menB*

ORF 043 shows 89.8% identity over a 128 aa overlap with a predicted ORF (ORF043.a) from

N. gonorrhoeae:

m043/g043

	10	20	30	40	50	60
m043.pep	MVVSNQNIYAAGPSALLHIRRQKSVMPSERFVEPSRVAVAAKVHGGLDGAAGFDEGERVF					
	: : : : : :					
g043	MVVSNQNIYAVGPSALFHIRRQKSVMPSERFVEPSRVAVAAKVHRGLDGAARFDEGERVF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m043.pep	QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFDGDQRTGEFVLQDVGGFVYAPTAVTV					
	: : : : : :					
g043	QPQAAQASGDGFAGLRFEIAFQVAFVQSDAAGDFDGDQRTAGEFAVQNIGGFVYAPAAVAV					
	70	80	90	100	110	120
	130					
m043.pep	VVAAEGEAQX					
g043	VVAAEGEAXX					
	130					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 179>:

```

a043.seq
1  ATGGTTGTTT CAAATCAAAA TATCTATGCC GCCGGCCCT CAGCACTTCT
51  TCACATCCGA AGGCAAAAAT CCGTAATGCC GTCTGAACGC TTCGTTGAAC
101 CGTCCCGCGT GCGGCTAGCC GCAAAAGTGC ATGGCGGCTT GGACGGTGCT

```

242

```

151 GCCGGATTCG ATGAGGGCGA GCGCGTGTT CAGCCGCAGG CGGCGCAGGC
201 ATCCGGCGAC GGTTCGCGG GTTTGCCTT TGAATAGCG TTCAGGTAG
251 CATTCTGTTCA GTCCGACGCG GCGGCGGATT TCGGCGATGG TCAGCGGACG
301 GGCGAATTCG TGTTGCAGGA TGTCGGCGGC TTCGTCTATG CGCCGACGGC
351 GGTAACCGTT GTCGTGGCGG CGGAAGGTGA AGCGCAATAA

```

This corresponds to the amino acid sequence <SEQ ID 180; ORF 043.a>:

```

a043.pep
1  MVVSNQNIYA AGPSALLHIR RQKSVMPSE FVEPSRVAVA AKVHGGLDGA
51  AGFDEGERVF QPQAAQASGD GFAGLRFEIA FQVAFVQSDA AGDFGDGQRT
101 GEFVLQDVGG FVYAPTAVTV VVAAEGEAQ*

```

m043/a043 100.0% identity in 129 aa overlap

	10	20	30	40	50	60
m043.pep	MVVSNQNIYAAGPSALLHIRRQKSVMPSE	FVEPSRVAVA	AKVHGGLDGA	AGFDEGERVF		
a043	MVVSNQNIYAAGPSALLHIRRQKSVMPSE	FVEPSRVAVA	AKVHGGLDGA	AGFDEGERVF		
	10	20	30	40	50	60
m043.pep	QPQAAQASGDGFAGLRFEIAFQVAFVQSDA	AGDFGDGQRT	GEFVLQDVGGFVYAPTAVTV			
a043	QPQAAQASGDGFAGLRFEIAFQVAFVQSDA	AGDFGDGQRT	GEFVLQDVGGFVYAPTAVTV			
	70	80	90	100	110	120
m043.pep	QPQAAQASGDGFAGLRFEIAFQVAFVQSDA	AGDFGDGQRT	GEFVLQDVGGFVYAPTAVTV			
a043	QPQAAQASGDGFAGLRFEIAFQVAFVQSDA	AGDFGDGQRT	GEFVLQDVGGFVYAPTAVTV			
	70	80	90	100	110	120
m043.pep	VVAAEGEAQX					
a043	VVAAEGEAQX					
	130					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 181>:

```

g044.seq
1  ATGCTGCCCC ACCAGAGCGT CGAGTTCTTG CCACAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GCGGCGCGTT TTCCAGCCGT TCGCTTCCA ACCGTGTATC
101 CAGTTTTTCCA TGCCGTTTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTACAGCC TTCGATAACG GCGGTCAGCT
201 CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 CGGCTGCCGT AGCGCATTA

```

This corresponds to the amino acid sequence <SEQ ID 182; ORF 044.ng>:

```

g044.pep
1  MLPDQSVEFL PQVVFDGLF GGGFPAVALP TVYPVFHAVF DVLRVGADDD
51  GAAAFERFQP FDNGGQLHAV VGGLRFAAEK FFFAAVAH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 183>:

```

m044.seq
1  ATGCCGTCCG ACTAGAGCGT CGAGTTCTTT CCAGAAGTCG TCGTTTTTGA
51  CGGGCTGTTT GGAGGCGGTT TTCCAGCCGT TCGCTTCCA ACCGTGTATC
101 CAGTTTTTCCA TGCCATTTT GACGTATTGC GAGTCGGTGC AGATGATGAC
151 GGTGCAGCGG CGTTTGAGCG ATTTAGTCC TTCGATGACG GCAGTCAGTT
201 CCATGCGGTT GTTGGTGGTT TGCGCTTCGC CGCCGAAAAG TTCTTTTTCG
251 TGGCTACCGT AGCGCAyTaa

```

This corresponds to the amino acid sequence <SEQ ID 184; ORF 044>:

```

m044.pep
1  MPSDXSVEFF PEVVFDGLF GGGFPAVALP TVYPVFHAI F DVLRVGADDD
51  GAAAFERFQS FDDGSQFHAV VGGLRFAAEK FFFVATVAH*

```


-- This corresponds to the amino acid sequence <SEQ ID 186; ORF 044.a>:

m044/a044 91.0% identity over a 89 aa overlap

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 044 shows 86.5% identity over a 89 aa overlap with a predicted ORF (ORF 044.ng) from *N. gonorrhoeae*:

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 187>:

```
g046.seq
1  ATGTCTGGCAA TGCTGCGTCC GACAAGCAGC CCGCCGCGcc gCGCCTGTAT
51  GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTCTCGGGG
151 CTGATGGTTT CGGTTATGCC gaATATGGAA AGGTACCGCt TTTcGTGTT
201 TTCGAGCTTG GGGCTGAGGT ATTCTAGGTA TtcgctGGAA CGACGCGCG
251 CGATCGCGCC GGGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGagaGCGCG AGcagcaagt cggcatcttc
351 Cgcgccggcg Cgttataatg tgAAGGGGGA TCGCcggttg ccgaAAACGG
401 TTTGGacatc gaggcggctg CTTGTTTCTT GCAATGCTTT TTCGTCGATG
451 TCGATAacgg TTACGTCGT CTGGTGATCG CGGGCAAGGT TTTGCGCGAG
```

501 GGTAGAACCT ACCTGCCCCG TGCCTAAAAT GAGGATTTTC ACGGTATGGG
551 TCGCCGGGTG A

This corresponds to the amino acid sequence <SEQ ID 188; ORF 046.ng>:

g046.pep
1 MSAMLRPTSS PPRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101 MLVSSLRESA SSKSASSAPA RYNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151 SITVTSLLVM AARFCATVEP TCPLPKMRIF TVWVAG*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 189>:

m046.seq
1 ATGTCGGCAA TGCTGCGTCC GACAAGCAst CCGC.r.sgc gCGcCTGTAT
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCCGGG
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTGTGT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
251 CGATGCGGCC GGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TCGCCCGTTG CCGAAAACGG
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
501 GG TAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 190; ORF 046>:

m046.pep
1 MSAMLRPTSX PXRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 191>:

a046.seq
1 ATGTCGGCAA TGCTGCGTCC GACAAGCAGT CCGCCGCGCC GCGCCTGTAT
51 GATGACCATC CGCACGCGGT CGTCTGCAAA ACGTAAAACC TGCAATGCGC
101 CCGGGCAGTC TATCAGGCCG GCAAGCTGTT CGGTAACGAG CTGTTCCGGG
151 CTGATGGTTT CGGTTATGCC GAATATGGAA AGGCTGCCGT TTTCGTGTGT
201 TTCGAGCTTG GGGCTGAGGT ATTCGAGGTA TTCGCTGGAA CGGACGCGCG
251 CGATGCGGCC GGGATGTTG AACAGGTCGG CGGCAACTTT GCAGGCGACG
301 ATGTTGGTTT CGTCGCTGCG GGAGAGCGCG AGCAGCAAGT CGGCATCTTC
351 CGCGCCGGCG CGTTCTAATG TGAAGGGGGA TCGCCCGTTG CCGAAAACGG
401 TTTGGACATC GAGGCGGCTG CCTGTTTCCT GCAATGCTTT TTCGTCGATG
451 TCGATAACGG TTACGTCGTT GTTGGGTATG GCGGCAAGGT TTTGTGCGAC
501 GG TAGAACCT ACCTGTCCGT TGCCTAAAAT GAGGATTTTC ACGGTGTGGG
551 TCGCCGAGTG A

This corresponds to the amino acid sequence <SEQ ID 192; ORF 046.a>:

a046.pep
1 MSAMLRPTSS PPRACMMTI RTRSSAKRKT CNAPGQSIRP ASCSVTSCSG
51 LMVSVMPNME RLPFSLFSSL GLRYSRYSLE RTRAMRPGML NRSAATLQAT
101 MLVSSLRESA SSKSASSAPA RSNVKGDAPL PKTVWTSRRL PVSCNAFSSM
151 SITVTSLLGM AARFCATVEP TCPLPKMRIF TVWVAE*

m046/a046 98.4% identity over a 186 aa overlap

	10	20	30	40	50	60
m046.pep	MSAMLRPTSPXRACMMTIRTRSSAKRKT CNAPGQSIRP ASCSVTSCSGLMVSVMPNME					
a046	MSAMLRPTSSPPRACMMTIRTRSSAKRKT CNAPGQSIRP ASCSVTSCSGLMVSVMPNME					

245

	10	20	30	40	50	60
m046.pep	70	80	90	100	110	120
	RLPFSLFSSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
a046	RLPFSLFSSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
	70	80	90	100	110	120
m046.pep	130	140	150	160	170	180
	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
a046	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
	130	140	150	160	170	180
m046.pep	TVWVAEX					
a046	TVWVAEX					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 046 shows 97.3% identity over a 185 aa overlap with a predicted ORF (ORF 046.ng) from *N. gonorrhoeae*:

m046/g046

	10	20	30	40	50	60
m046.pep	MSAMLRPTSPXPRACMMTIRTRSSAKRKTCNAPGQSIRPASCSTVSCSGLMVSVMPNME					
g046	MSAMLRPTSSPFRACMMTIRTRSSAKRKTCNAPGQSIRPASCSTVSCSGLMVSVMPNME					
	10	20	30	40	50	60
m046.pep	70	80	90	100	110	120
	RLPFSLFSSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
g046	RLPFSLFSSSLGLRYSRYSLERTRAMRPGMLNRSAAATLQATMLVSSLRESASSKSASSAPA					
	70	80	90	100	110	120
m046.pep	130	140	150	160	170	180
	RSNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLGMAARFCATVEPTCPLPKMRIF					
g046	RYNVKGDAPLPKTVWTSRRLPVSCNAFSSMSITVTSLLVMAARFCATVEPTCPLPKMRIF					
	130	140	150	160	170	180
m046.pep	TVWVAEX					
g046	TVWVAGX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 193>:

g047.seq

```

1  ATGGTCATCA TACAGGCGcg gcGCGGCGGG CTGCTGTGCG GACGCAGCAT
51 TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CCGCGCCGCA AACCGTCATC
151 ATCGAAGGCG ACGAAATCCT GTTTGCCGCC GCCGCCGAAA ACATCGGGGC
201 GGTCATACcC gaATTGCGCC CCAAAGAAAC CAGCACCCGC CGCATCATGA
251 TGCCGCGCGG CGGCAACATc tgctACCGCC TCGCCAAGCA GCTCGAACAC

```

246

```

301 GCATAcaacG TCAAAATCAT CGAATGCCGG CCGCGCcggtg ccgaATGGAT
351 AGCCGAAAAc ctcgAcaaCA CCCTCGTCCT GCAAGGTTTCG Gcaaccgacg
401 aAaccctgct cgAcaacgaa tacatcgacg aaatcgaCGT ATTCTGCGCC
451 CTGACCAACG ACGACGAAAG CAACATTAtg tCCGCCCTTT TGGCGAAAAA
501 CCTcggcgCG AAGCgcgtca tcggCATCGT CAACCGCTCA AGCTACGTCG
551 ATTTGCTCGA AGGCAACAAA ATCGACATCG TCGTCTCCCC CCACCTCATC
601 ACCATCGGCT CGATACTCGC CCACATCCGG CGCGGCGACA TCGTTGCCGT
651 CCACCCCATC CGGCGCGGCA CGGCGGAAGC CATCGAAGTC GTCGCGCAGC
701 GCGACAAAAA AACTTCGCCC ATCATCGGCA GGCGCATCAG CGGCATCAA
751 TGGCCCGAAG GCTGCCACAT TGCCGCCGTC GTCCGCGCCG GAACCGGCGA
801 AACCATTATG GGACACCATA CCGAAACCGT CATCCAAGAC GGTGACCACA
851 TCATCTTTTT CGTCTCGCGC CGGCGCATCC TGAACGAACT GGAGAAACTC
901 ATCCAAGTCA AAATGGGCTT TTTCCGATAA

```

This corresponds to the amino acid sequence <SEQ ID 194; ORF 047.ng>:

g047.pep

```

1 MVIIQARRGG LLVGRSIADI AQLDPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI CYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDESNI SALLAKNLGA KRVIGIVNRS SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTS A IIGRRISGIK
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 195>:

m047.seq

```

1 ATGGTCATCA TACAGgCGcG C..syGCGGA sTGCTTGTCG GACGCAGCAT
51 TGCCGACATC GCCCAAGATT TGCCCGACGG GGCCGACTGC CAAATCTGCG
101 CCGTTTACCG CAACAACCGC CTCATCGTCC CGCGCGCGCA AACCGTCATC
151 ATCGAAGGCG ACGAAATCCT ATTGCGGCC GCGCCGAAA ACATCGGCGC
201 GGTCAATACC GAATGCGCC CCAAAGAAAC CCAAAGAAAC CAGCcCmgmm
251 GcATCATGAT TkCCGGCGGC GGCAACATCG GTACCGTCT CGCCAAGCAG
301 CTCGAACACG CATACAACGT yAAAAATCATC GAATGCCGGC CGCGCCGTGC
351 CGAATGGATA GCCGAAAACC TCGACAACAC CCTCGTCyTG CAAGGTTTCG
401 CAACCGACGA AACCTGCTC GACAACGAAT ACATCGACGA AATCGACGTA
451 TTCTGCGCCC TGACCAACGA CGACGAAAGC AACATTATGT CCGCCCTTTT
501 GGCGAaAAAC CTCGGCGCGA AGCGCGTCAT CGGCATCGTC AACCGCTCAA
551 GCTACGTCGA TTTGCTCGAA GGCAACAAAA TCGACATCGT CGTCTCCCCC
601 CACCTCATCA CCATCGGCTC GATACTCGCC CACATCCGGC GCGGCGACAT
651 CGTTGCCGTC CACCCATCC GGCGCGGCAC GGCGGAAGCC ATCGAAGTCG
701 TCGCACACGG CGACAAAAA ACTTCCGCCA TCATCGGCAG GCGCATCAGC
751 GGCAATCAAAT GGCCCGAAGG CTGCCACATT GCCGCCGTCG TCCGCGCCGG
801 AACCGGCGAA ACCATTATGG GACACCATAC CGAAACCGTC ATCCAAGACG
851 GCGACCACAT CATCTTTTTC GTCTCGCGCC GGCGCATCCT GAACGAACTG
901 GAAAAACTCA TCCAGGTCAA AATGGGCTTT TTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 196; ORF 047>:

m047.pep

```

1 MVIIQARXXG XLVGRSIADI AQLDPDGADC QICAVYRNNR LIVPAPQTVI
51 IEGDEILFAA AAENIGAVIP ELRPKETORN QPXXIMIXGG GNIGYRLAQ
101 LEHAYNVKII ECRPRRAEWI AENLDNTLV LQGSATDETL DNEYIDEIDV
151 FCALTNDDES NIMSALLAKN LGA KRIVIGIV NRSSYVDLLE GNKIDIVVSP
201 HLITIGSILA HIRRGDIVAV HPIRRGTAEA IEVVAHGDKK TSAIIGRRIS
251 GIKWPEGCHI AAVVRAGTGE TIMGHHTETV IQDGDHIIFF VSRRRIINEL
301 EKLIQVKMGF FG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 197>:

a047.seq

```

1 ATGGTCATCA TACAGGCGCG GCGCGGCGGA CTGCTTGTCG GACGCAGCAT

```

247

```

51  TGCCGACATC  GCCCAAGATT  TGCCCGACGG  GGCCGACTGC  CAAATCTGCG
101 CCGTTTACCG  CAACAACCGC  CTCATCGTCC  CCGCGCCGCA  AACCCTCATC
151 ATCGAAGGCG  ACGAAATCCT  ATTTGCCGCC  GCCGCCGAAA  ACATCGGCGC
201 GGTACATACC  GAATTGCGCC  CCAAAGAAAC  CAGCACCCGC  CGCATCATGA
251 TTGCCGGCGG  CGGCAACATC  GGCTACCGTC  TCGCCAAGCA  GCTCGAACAC
301 GCATACAACG  TCAAATCAT  CGAATGCCGG  CCGCGCCGTG  CCGAATGGAT
351 AGCCGAAAAC  CTCGACAACA  CCCTCGTCCT  GCAAGGTTTC  GCAACCGACG
401 AAACCTTGCT  CGACAACGAA  TACATCGACG  AAATCGACGT  ATTCTGCGCC
451 CTGACCAACG  ACGACGAAAG  CAACATTATG  TCCGCCCTTT  TGGCGAAAAA
501 CCTCGGCGCG  AAGCGCGTCA  TCGGCATCGT  CAACCGCTCA  AGCTACGTCG
551 ATTTGCTCGA  AGGCAACAAA  ATCGACATCG  TCGTCTCCCC  CCACCTCATC
601 ACCATCGGCT  CGATACTCGC  CCACATCCGG  CGCGGCGACA  TCGTTGCGGT
651 CCACCCATC  CGGCGCGGCA  CGGCGGAAGC  CATCGAAGTC  GTCGCACACG
701 GCGACAAAAA  AACTTCCGCC  ATCATCGGCA  GGCGCATCAG  CGGCATCAAA
751 TGGCCCGAAG  GCTGCCACAT  TGCCGCGGTC  GTCCGCGCCG  GAACCGGCGA
801 TIGSILAHIR  RGDIVAVHPI  RRGTAEAIEV  VAHGDKKTS  IIGRRISGKI
851 TCATCTTTT  CGTCTCGCGC  CGGCGCATCC  TGAACGAACT  GGAAAACTC
901 ATCCAAGTCA  AAATGGGCTT  TTTCGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 198; ORF 047.a>:

```

a047.pep
1  MVIIQARRGG LLVGRSIADI AQDLDPGADC QICAVYRNNR LIVPAPQTVI
51  IEGDEILFAA AAENIGAVIP ELRPKETSTR RIMIAGGGNI GYRLAKQLEH
101 AYNVKIIECR PRRAEWIAEN LDNTLVLQGS ATDETLLDNE YIDEIDVFCA
151 LTNDDESNI SALLAKNLGA KRIVIGVNR SYVDLLEGNK IDIVVSPHLI
201 TIGSILAHIR RGDIVAVHPI RRGTAEAIEV VAHGDKKTS IIGRRISGKI
251 WPEGCHIAAV VRAGTGETIM GHHTETVIQD GDHIIFFVSR RRILNELEKL
301 IQVKMGFFG*

```

m047/a047 96.5% identity over a 312 aa overlap

```

              10      20      30      40      50      60
m047.pep  MVIIQARXXGLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIIEGDEILFAA
          ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a047      MVIIQARRGGLLVGRSIADIAQDLDPGADCQICAVYRNNRLIVPAPQTVIIIEGDEILFAA
              10      20      30      40      50      60

              70      80      90     100     110     120
m047.pep  AAENIGAVIPELRPKETQRNPXXIMIXGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
          ||||| ||||| : ||| ||||| ||||| ||||| ||||| ||||| |||||
a047      AAENIGAVIPELRPKETSTR---IMIAGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI
              70      80      90     100     110

              130     140     150     160     170     180
m047.pep  AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRIVIGIV
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a047      AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDESNIMSALLAKNLGAKRIVIGIV
              120     130     140     150     160     170

              190     200     210     220     230     240
m047.pep  NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a047      NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK
              180     190     200     210     220     230

              250     260     270     280     290     300
m047.pep  TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a047      TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQDGDHIIFFVSRRRILNEL
              240     250     260     270     280     290

              310
m047.pep  EKLIQVKMGFFGX
          ||||| |||||

```

248

a047 EKLIQVKMGFFGX
 300 310

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 047 shows 96.2% identity over a 312 aa overlap with a predicted ORF (ORF 047.ng)
from *N. gonorrhoeae*:

m047/g045

m047.pep	MVIIQARXXGXLVGRSIADIAQDLDPGADCQICAVYRNRLIVPAPQTVIIIEGDEILFAA	60
g047	MVIIQARRGGLLVGRSIADIAQDLDPGADCQICAVYRNRLIVPAPQTVIIIEGDEILFAA	60
m047.pep	AAENIGAVIPELRPKETQRNQXXIMIXGGGNIGYRLAKQLEHAYNVKIIIECRPRRAEWI	120
g047	AAENIGAVIPELRPKETSTR---IMIAGGGNICYRLAKQLEHAYNVKIIIECRPRRAEWI	117
m047.pep	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDES NIMSALLAKNLGAKRIVIGIV	180
g047	AENLDNTLVLQGSATDETLLDNEYIDEIDVFCALTNDDES NIMSALLAKNLGAKRIVIGIV	177
m047.pep	NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	240
g047	NRSSYVDLLEGNKIDIVVSPHLITIGSILAHIRRGDIVAVHPIRRGTAEAEVVAHGDKK	237
m047.pep	TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQGDHIIFFVSRRRILNEL	300
g047	TSAIIGRRISGIKWPEGCHIAAVVRAGTGETIMGHHTETVIQGDHIIFFVSRRRILNEL	297
m047.pep	EKLIQVKMGFFGX 313	
g047	EKLIQVKMGFFGX 310	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 199>:

g048.seq

1	ATGCTCGACA AAGGCGAGGA GTTGCCCGTC GATTTCACCA ACCGCCTGAT
51	TTACTACGTc ggcCCcgTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCCG
101	CAGGTCCGAC CACAGCCACC CGCATGGACA AATTTACCCG CCAAATGCTC
151	AAACAAACCG GCCTCTTGGG CATGATCGGC AAATCCGagc gcgGcgcggc
201	cacctGCGAA GCcatCGCCG ACAACAAGGC CGTGACCTC ATGGCAGTCG
251	GCGGCGCGGC ATACCTCGTG GCAAAAGCCA TCAAATCTC CAAAGTCTTG
301	GCGTTCCCCG AATTGGGTAT GGAAGCCGTT TACGAATTG AAGTCAAAGA
351	TATGCCCGTA ACCGTCGCCG TGGACAGCAA AGGCGAATCC ATCCACGCCA
401	CCGCCCGCG CAAATGGCAG GCGAAAATCG GCATCATCCC CGTCGAGTCT
451	TGA

This corresponds to the amino acid sequence <SEQ ID 200; ORF 048.ng>:

g048.pep

1	MLDKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTROML
51	KQTGLLG MIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101	AFPELGMEAV YEFVCKMPV TVAVDSKES IHATAPRKWQ AKIGIIPVES
151	*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 201>:

m048.seq

1	ATGCTCAACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
51	TTACTACGTC GGCCCGTCG ATCCGGTCGG CGATGAAGTC GTCGGTCCCG

```
m048.pep
  1 MLNKGEELPV DFTNRLIYYV GPVDPVGDEV VGPAGPTTAT RMDKFTRQML
51 EQTDLLGMIG KSERGVATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 APFELGMEAI YEFVKKDMPV TVAVDSKGES IHATAPRKWQ AKIGIIPVES
151 *
```

```
a048.seq
1 ATGCTCGACA AAGGCGAAGA ATTGCCCGTC GATTTCACCA ACCGCCTGAT
51 TTA CTACGTC GCGCCCGTCG ATCCGGTCGG CGACGAAATC GTCGGCCAG
101 CAGGTCGAC ACCGCCACC CGCATGGACA AATTCACCCG CCAATGCTC
151 GAACAAACCG ACCTCTTGGG CATGATCGGC AAATCCGAGC GCGCGCGGG
201 CACCTCGGAA GCCATCGCCG ACAACAAAGC CGTGACTCTC ATGGCATCTG
251 GCGGCGCGGC GTATCTCGTG GCAAAAGCCA TCAAATCTTC CAAAGTCTTG
301 GCGTTCCCGC AATTGGGCAT GGAAGCCATT TACGAATTG AAGTCAAAGA
351 CATGCCCGTA ACCGTCGCCG TAGACAGCAA AGGCGAATCC ATCCACGCCA
401 CGGCCCGGCC CCAATGGCAG GCGAAAATCG GCATCATCCC CGTCAAATCT
451 TGA
```

```
a048.pep
1  MLDKGEELPV DFTNRLIYV GPVDPVGDEI VGPAGPTTAT RMDKFTRQML
51  EQTDLLGMIG KSERGAATCE AIADNKAVYL MAVGGAAYLV AKAIKSSKVL
101 AFPELGMEAI YEFEVKDMPV TVAVDSKGES IHATAPPQWQ AKIGIIPVKS
151 *
```

	10	20	30	40	50	60
m048.pep	MLNKGEELPVDF	TNRLIIYYGV	FPVDPVGDE	VGPAGPTTAT	RMDFTRQMLE	QTDLLGMIG
	: : : : : :	: : : : : :	: : : : : :	: : : : : :	: : : : : :	: : : : : :
a048	MLDKGEELPVDF	TNRLIIYYGV	FPVDPVGDE	IVGPAGPTTAT	RMDFTRQMLE	QTDLLGMIG
	10	20	30	40	50	60
	70	80	90	100	110	120
m048.pep	KSERGVATCEAI	ADNKAVIYLM	AVGGAAYLV	AKAIKSSKVL	AFPELGMEAI	YEFEVKDMFV
	: : : : : :	: : : : : :	: : : : : :	: : : : : :	: : : : : :	: : : : : :
a048	KSERGAATCEAI	ADNKAVIYLM	AVGGAAYLV	AKAIKSSKVL	AFPELGMEAI	YEFEVKDMFV
	70	80	90	100	110	120
	130	140	150			
m048.pep	TVAVDSKGESIH	ATAPRKWQAK	IGIIPVESX			
	: : : : : :	: : : : : :	: : : : : :			
a048	TVAVDSKGESIH	ATAPPQWQAK	IGIIPVKXS			
	130	140	150			

ORF 048 shows 96.4% identity over a 150 aa overlap with a predicted ORF (ORF 048.ng) from *N. gonorrhoeae*:

250

m048/g048

	10	20	30	40	50	60
m048.pep	MLNKGEELPVDFTNRLIYYVGPVDPVGVGDEVVGPAGPTTATRMDKFTRQMLEQTDLLGMIG					
	:					
g048	MLDKGEELPVDFTNRLIYYVGPVDPVGVGDEVVGPAGPTTATRMDKFTRQMLKQTGLLGMIG					
	10	20	30	40	50	60

	70	80	90	100	110	120
m048.pep	KSERGVATCEAIADNKAVYLMVAVGGAAYLVAKAIKSSKVLAFPELGMEAIYEFVVKDMPV					
	:					
g048	KSERGAATCEAIADNKAVYLMVAVGGAAYLVAKAIKSSKVLAFPELGMEAVYEFVVKDMPV					
	70	80	90	100	110	120

	130	140	150
m048.pep	TVAVDSKGESIHATAPRKWQAKIGIIPVESX		
g048	TVAVDSKGESIHATAPRKWQAKIGIIPVESX		
	130	140	150

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 205>:

g049.seq

1	ATGCGGGCGC	AGGCGTTTGA	TCAACCGTTC	GGTCAGCTCC	TGTTCCGACA
51	GGCAGAACAC	TTTCGCGCCG	TTGACGGCTT	TCGGGTTTCA	GATATTGATT
101	TGGACGGGCA	TCAACGCCTC	TTCCGCACCG	CCTTCGCCGT	TTTCCGCAAC
151	CCCGTCTGCC	GCCGTACCGG	ATTCTGCCGC	ATCGGCGTTT	TCCCCGCCCT
201	CAATCTGTGC	GGTTTCAAAT	TCGGCACTGT	CTTTTTTGGC	ATCGAACCAG
251	ATTCTCCGCC	GCGATTCGAT	GTGTTTTTCC	GAAAccggca	tTTGCAGGGA
301	AGCCTgcgcg	TTGAGCCAGT	TTCTCTGAAG	GACGATCATC	GGGTCGGTTT
351	CGACTTCCTC	GCCGCAATCG	GCAACGGCgc	tGTTGTGTTC	TTCTGCCAT
401	TTCTTCAGAT	ACGCCTTTAA			

This corresponds to the amino acid sequence <SEQ ID 206; ORF 049.ng>:

g049.pep

1	MRAQAFDQPF	GQLLFGQAEH	FAPVDGFRVQ	DIDLGHQRL	PRTAFVFRN
51	PVCRRTGFCR	IGVFPALNLC	GFKFGTVFFG	IEPDSPPRFD	VFFRNRHLQG
101	SLRVEPVFLK	DDHRVGFDFL	AAIGNGAVVF	FLPFLQIRL*	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 207>:

m049.seq (partial)

1	ATGCGGGCGC	AGGCGTTTGA	TCAGCCGTTC	GGTCAGCTCC	TGTTCCGACA
51	GGCAGAACAC	TTTCGCGCCG	TTGACGGCTT	TCGGGTTTCA	GATATTGATT
101	TGGACGGGCA	TCAACGTTTC	TTCCGCATCG	TTTCCCCCGT	TTTCCGAAAC
151	CGCCGGCTCA	TTCGTGCCCG	ATTCTGCCTC	GTCGGCGTTT	TCCCCGCTTT
201	CAATCTGTCC	GGTTTCAAAT	TCGACACTGT	CTTTTTTGGT	ATCAAACCGG
251	ATTCTCCGCC	GCGATTCGAT	GTGTTTTTCC	GAAACCGACA	TTTGCAGGGA
301	AGCCTGCGCG	TTGAGCCAGT	TTCTCTGAAG	GACGATCATC	GGGTCGGTTT
351	CGACTTCCTC	GCCGCAATCG	GCAACGGCGG	CATTGTGTTC	CTCCTGCCAT
401	TTTTTCAGAT	ACGCCTT...			

This corresponds to the amino acid sequence <SEQ ID 208; ORF 049>:

m049.pep (partial)

1	MRAQAFDQPF	GQLLFGQAEH	FAPVDGFRVQ	DIDLGHQRF	FRIVFPVFRN
51	RRLIRAGFCL	VGVPFANLS	GFKFDTVFFG	IKPDSPPRFD	VFFRNRHLQG
101	SLRVEPVFLK	DDHRVGFDFL	AAIGNGGIVF	LLPFLQIRL...	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 209>:

a049.seq

a049.pcp

1 MRAQAFDQPF GQLLFGQAEH FAPVDGFRVQ NIDLGDGHQRF FRTAFAVFRN
51 PVCRRTRFCR IGVFPAFNLS GFKEGTVFFG IKPDSPPRFD VFFRNRHLQG
101 SLRVEPVFLK DDHRVGFDL AAINGGGIVF LLPFFQIRL

		10	20	30	40	50	60
m049.pep		MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQDIDL	DGHQRRFFRIVFPVFRNRRLIRAGFCL				
a049		MRAQAFDQPFQQLLFGQAEHFAPVDGFRVQ	NDL	DGHQRRFFRTAF	AVFRNPVCRRTRECR		
		10	20	30	40	50	60
		70	80	90	100	110	120
m049.pep		VGVFPAFNL	SGFKFDTVFFG	IKPDSPPRFDVFFRNRHLQGS	LRVEPVFLKDDHRV	GFDFL	
a049		IGVFPAFNL	SGFKFTVFFG	IKPDSPPRFDVFFRNRHLQGS	LRVEPVFLKDDHRV	GFDFL	
		70	80	90	100	110	120
		130	139				
m049.pep		AAIGNGGIV	FLLPFFQIRL				
a049		AAIGNGGIV	FLLPFFQIRL				
		130					

Homology with a predicted ORF from *N. gonorrhoeae*

m049/q049

		10	20	30	40	50	60
m049.pep		MRAQAFDQPFQGQLLF	GQAEHFAPVDGFRVQ	IDLDGHQRFFRIVFP	VFVRNRLIRAGFCL		
g049		MRAQAFDQPFQGQLLF	GQAEHFAPVDGFRVQ	IDLDGHQRLFR	TAFVFRNPFVCRRT	GF	CF
		10	20	30	40	50	60
		70	80	90	100	110	120
m049.pep		VGVFPAFNLSGFKFD	TVFFGIKPDSPPRF	DVFFRNRHLQGS	LRVEPVFLKDDH	RVGFDFL	
g049		IGVFPAFNLSGFKF	GTVFFGIKPDSPPR	FVFFRNRHLQGS	LRVEPVFLKDDH	RVGFDFL	
		70	80	90	100	110	120
		130	139				
m049.pep		AAIGNGGIVFLLP	FFQIRL				
g049		AAIGNGAVVFLP	FLQIRLX				
		130	140				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 211>:

```
g050.seq
1  atgggCGCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGg
51  cagccccGAA AAAGccgtgt TGATGGcaaa AGAATCCCTG ATGAGCCACA
101 TCGAcatCca aGaATTGCAG GAAAAAGCCG CGTccggggc ggaattgtcc
151 accaccgaAG ccCTGCGCCT cGAACTCTTT GAAAAAGGTCA ACGCGCTGGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAT GCGGCCTCCA AACCGATTGC CATGATTCCC
301 AACTGTGCcg ccacCCGcca cgtcgAATTT GAATTGgACG GCTCAGGtcc
351 TGTCGAactc acgccGCcgc gtgtCGAAGA CTGA
```

This corresponds to the amino acid sequence <SEQ ID 212; ORF 050.ng>:

```
g050.pep
1  MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVED*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 213>:

```
m050.seq
1  ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGTATCG GCATCGGCGG
51  C..agCCgAA AAAGCCGTGC TGATGGCAAA AGAGTCCCTG ATGAGCCACA
101 TCGACATTCA AGAATTGCAG GAAAAGGCCG CGTCCGGCGC GgAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTGGG
201 CATCGGCGCA CAAGGCTTGG GCGGACTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA TCCGACCCAC GCGGCCTCCA AACCGATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351 TGTCGAACTC ACGCCGCCGC GCGTCGAAGA TGGCCCGATT TGA
```

This corresponds to the amino acid sequence <SEQ ID 214; ORF 050>:

```
m050.pep
1  MGAGWCPPGI LGIGIGGXAE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDGPI *
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 215>:

```
a050.seq
1  ATGGGCGCGG GCTGGTGTCC TCCCGGCATC TTGGGCATCG GCATCGGCGG
51  TACGCCCGAA AAAGCCGTGT TGATGGCGAA AGAATCCCTG ATGAGCCACA
101 TCGACATCCA AGAATTGCAG GAAAAGGCCG CGTCCGGCGC GGAATTGTCC
151 ACCACCGAAG CCCTGCGCCT CGAACTCTTT GAAAAAGTCA ACGCGCTAGG
201 CATCGGCGCG CAAGGCTTGG GCGGTCTGAC CACCGTGTG GACGTGAAAA
251 TCCTCGATTA CCCGACCCAC GCGGCCTCCA AACCGATTGC CATGATTCCG
301 AACTGCGCCG CCACCCGCCA CGTCGAATTT GAATTGGACG GCTCAGGCCC
351 TGTCGAACTC ACGCCGCCGC GCGTCGAAGA CTGGCCC
```

This corresponds to the amino acid sequence <SEQ ID 216; ORF 050.a>:

```
a050.pep
1  MGAGWCPPGI LGIGIGGTPE KAVLMAKESL MSHIDIQELQ EKAASGAELS
51  TTEALRLELF EKVNALGIGA QGLGGLTTVL DVKILDYPH AASKPIAMIP
101 NCAATRHVEF ELDGSGPVEL TPRVEDWP
```

m050/a050 97.7% identity over a 129 aa overlap

```

          10      20      30      40      50      60
m050.pep  MGAGWCPPGILGIGIGGXAEKAVLMAKESLSMSHIDIQELQEKAASGAELSTTEALRLELF
          |||||: |||||
a050      MGAGWCPPGILGIGIGGTPEKAVLMAKESLSMSHIDIQELQEKAASGAELSTTEALRLELF
```

253

	10	20	30	40	50	60
m050.pep	70	80	90	100	110	120
	EKNVALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
a050	EKNVALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
a050	TPPRVEDWP					

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 050 shows 98.4% identity over a 127 aa overlap with a predicted ORF (ORF 050.ng) from *N. gonorrhoeae*:

m050/g050

	10	20	30	40	50	60
m050.pep	70	80	90	100	110	120
	MGAGWCPPGILGIGIGGXAEKAVLMAKESLSMHSIDIQELQEKAASGAELSTTEALRLELF					
g050	MGAGWCPPGILGIGIGGTPEKAVLMAKESLSMHSIDIQELQEKAASGAELSTTEALRLELF					
	10	20	30	40	50	60
m050.pep	70	80	90	100	110	120
	EKNVALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
g050	EKNVALGIGAQGLGGLTTVLDVKILDYPHTHAASKPIAMIPNCAATRHVEFELDGSGPVEL					
	70	80	90	100	110	120
	130					
m050.pep	TPPRVEDGPIX					
g050	TPPRVEDX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 217>:

g050-1.seq

```

1  ATGACCGTTA TCAAGCAAGA AGACTTTATT CAAAGTATCT GCGATGCCTT
51  CCAATTCATC AGCTACTACC ATCCAAAGA CTACATCGAC GCGCTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCCGCCG CCAAAGACGC GATGACGCG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCCGAA AACAAACGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCCTCAA AGTCGGTATG GATGTGCAAT
251 GGGATGCGGA CATGAGCGTG GAAAAGATGG TTAACGAAGG CGTACGCCGC
301 GCCTACACTT GGAAGGCAA CACCCTGCGC GCTTCCGTCC TCGCCGATCC
351 GCGCGGCAAA CGCCAAAACA CCAAAGACAA CACCCCGGCC GTCATCCACA
401 TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GCGGCTCTG AAAACAAATC CAACTCGCT ATGCTCAACC CTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACGATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGC ATCGGCATCG GCGGCACgC CGAAAAAGCC
601 GTGTGTATGG cgaAAGAATC CCTGATGAGC CACATCGACA TCCAAGAATT
651 GCAGGAAAAA GCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAACT CTTTGAAGAAG GTCAACGCGC TGGGCATCGG CGCGCAAGGC
751 TTGGCGGTC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTACCCGAC
801 CCATGCCGCC TCCAAACCGA TTGCCATGAT TCCCAACTGT GCCGCCACCC
851 GCCACGTCGA ATTGAATTG GACGGCTCAG GTCCTGTCGA ACTCACGCCG
901 CCGCGCGTCG AAGACTGACC CGATCTGACT TACAGCCCGC ACAACGGCAA
951 ACGCGTCGAT GTCGATAAGC TGACCAAAGA AGAAGTGGCA AGCTGGAAAA
1001 CCGCGACGCT ATTGCTGTTG AACGGCAAAA TCCTCACCGG CCGCGATGCC
1051 GCGCACAAAC GCCTCGTCAA TATGCTCGAC AAAGGCGAGG AGTTGCCCGT
1101 CGATTTCAAC AACCGCTGTA TTTACTACGT CCGCCCGGTC GATCCGGTGC
1151 CCGATGAAGT CGTCGGTCCC GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTTACCC GCCAAATGCT CAAACAAACC GGCCTCTTGG GCATGATCGG

```

254

```

1251 CAAATCCGAG CGCGGCGCGG CCACCTGCGA AGCCATCGCC GACAACAAGG
1301 CCGTGTACCT CATGGCAGTC GCGGCGCGCG CATACCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGGTTCCCG GAATTGGGTA TGGAGCCGT
1401 TTACGAATTT GAAGTCAAAG ATATGCCCGT AACCGTCGCC GTGGACAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCGC GCAAATGGCA GCGGAAAATC
1501 GGCATCATCC CCGTCGAGTC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 218; ORF 050-1.ng>:

g050-1.pap

```

1  MTVIKQEDFI  QSICDAFQFI  SYHPKDYID  ALYKAWQKEE  NPAAKDAMTQ
51  ILVNSRMCAE  NNRPICQDTG  IATVFLKVGM  DVQWDADMSV  EKMVNEGVRR
101 AYTWEGNTLR  ASVLADPAGK  RQNTKDNTPA  VIHMSIVPGG  KVEVTCAAKG
151 GGSENKSKLA  MLNPSDNIVD  WVLKTIPTMG  AGWCPPGILG  IGIGGTPEKA
201 VLMAKESLMS  HIDIQELQEK  AASGAELSTT  EALRLELFEX  VNALGIGAQG
251 LGLLTTVLVD  KILDYPTTHA  SKPIAMIPNC  AATRHVEFEL  DSGSPVELTP
301 PRVED*PDLT  YSPDNGKRVD  VDKLTKEEVA  SWKTGDVLLL  NGKILTGRDA
351 AHKRLVNMLD  KGEELPVDFE  NRIIYVGVVP  DPGDEVVGP  AGPTTATRMD
401 KFTRQMLKQT  GLLGMIGKSE  RGAATCEAIA  DNKAVYLMV  GGAAYLVAKA
451 IKSSKVLAF  ELGMEAVYEF  EVKDMPTVTA  VDSKGESIHA  TAPRKWQAKI
501 GIIPVES*

```

g050-1/p14407

```

sp|P14407|FUMB_ECOLI FUMARATE HYDRATASE CLASS I, ANAEROBIC (FUMARASE)
>gi|280063|pir|B44511 fumarate hydratase (EC 4.2.1.2) fumb, iron-dependent - Escherichia coli
>gi|146048 (M27058) anaerobic class I fumarase (EC 4.2.1.2) [Escherichia coli] Length = 548
Score = 172 bits (432), Expect = 4e-42
Identities = 138/488 (28%), Positives = 216/488 (43%), Gaps = 22/488 (4%)

```

```

Query: 11  QSICDAFQFISYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAENNRPICQDTG 70
          Q+ DA + H K L+ E + K Q L NS + A+ P CQDTG
Sbjct: 53  QAFHDASFMLRPAHQKQVAAILHDPEASEND---KYVALQFLRNSEIAAKGVLPCTQDTG 109

```

```

Query: 71  IATVFLKVGMVDVQWDADMSVEKMVNEGVRRAYTWEGNTLRASVLADPAGKRQNTKDNTPA 130
          A + K G V W E+ +++GV Y E N + A K NT N PA
Sbjct: 110 TAIIVGKKQQRV-WTGGGD-EETLSKGVYNTYI-EDNLRYSQNAALDMYKEVNTGTNLPA 166

```

```

Query: 131 VIHMSIVPGGKVEVTCAAKGGGSENKSKL-----AMLNPSDNIVDWLKTIPMGAGWCP 185
          I + V G + + C AKGGG NK+ L A+L P + +++++ + T+G CP
Sbjct: 167 QIDLAVDGDGEYKFLCVAKGGGSANKTYLYQETKALLTPG-KLKNFLVERMRTLGTAAFP 225

```

```

Query: 186 PXXXXXXXXXTPEKAVLMAKESLSHIDIQELQEAASGAELSTTEALRLELFERVNXXX 245
          P T + L + +H EL + + L EL E+
Sbjct: 226 PYHIAFVIGGTSAETNLKTVKLASAHY-YDELPTEGNEHGQAFRDVQLEQELLEEAQKLG 284

```

```

Query: 246 XXXXXXXXXXXTTVLVDVKILDYPTTHAASKPIAMIPNCAATRHVEFELDGSF----PVELTPP 301
          D++++ P H AS P+ M +C+A R+++ +++ G +E P
Sbjct: 285 LGAQFGGKYFAH-DIRVIRLPRHGASCPVGMGVSADRNIAKINREGIWIEKLEHNPF 343

```

```

Query: 302 RVEDXPDLTYSYSPDNGKRVVDVKLTKE---EASWKTGDVLLNGKILTGRDAHKLRLVM 358
          + +VD+++ KE +++ + L L G I+ GRD AH +L +
Sbjct: 344 QYIPQELRQAGEGEAVKVDLNRPMKEILAQLSQYPVSTRSLTGTIIIVGRDIAHAKLKL 403

```

```

Query: 359 LDKGEELPVDFETNRIIYXXXXXXXXXXXXXXXXTTATRMDKFTRQMLKQTGLGMIGK 418
          +D G+ELP + IYY TTA RMD + + G + M+ K
Sbjct: 404 IDAGKELPQYIKDHPYIYAGPAKTPAGYPSGSLGPTTAGRMDSYVDLLQSHGGSMIMLAK 463

```

```

Query: 419 SERGAATCEAIADNKAVYLMVGG-AAYLVAKAIKSSKVLAFPELGMEAVYEFVVDMPV 477
          R +A + YL ++GG AA L ++IK + +A+PELGMEA+++ EV+D P
Sbjct: 464 GNRSQQVTDACHKHGGFYLGSIIGPAAVLAQSIKHLECVAYPELGMEAIWKIEVEDFPA 523

```

```

Query: 478 TVAVDSKG 485
          + VD KG
Sbjct: 524 FILVDKKG 531

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 219>:

m050-1.seq

```

1  ATGACCGTCA TCAAACAGGA AGACTTTATC CAAAGCATTT GCGATGCCTT
51  CCAATCATC AGCTACTATC ATCCCAAAGA CTACATCGAC CGCTTTTATA
101 AGGCGTGGCA GAAGGAAGAA AATCCTGCCG CCAAAGACGC GATGACGCAG
151 ATTTTGGTCA ACAGCCGTAT GTGTGCGGAA AACACCGGCC CCATCTGCCA
201 AGACACAGGT ATCGCAACCG TCTTCTCTCA AGTCGGTATG AACGTCCAAT
251 GGGATGCGGA CATGAGCGTG GAAGAGATGG TTAACGAAGG CGTACGCCGC

```

```

301 GCCTACACTT GGAAGGCAA TACGCTGCGC GCTTCCGTCC TCGCCGATCC
351 GGCCGGCAAA CGCCAAAACA CCAAGACAA CACCCCGGCC GTCATCCATA
401 TGAGCATCGT GCCGGGCGGT AAAGTCGAAG TAACCTGCGC GGCAAAAGGC
451 GGCGGCTCTG AAAACAAATC CAAACTCGCC ATGCTCAATC CTTCCGACAA
501 CATCGTCGAT TGGGTATTGA AAACCATCCC GACCATGGGC GCGGGCTGGT
551 GTCCTCCCGG CATCTTGGGT ATCGGCATCG GCGGCACGCC CGAAAAAGCC
601 GTGCTGATGG CAAAGAGTCT CCTGATGAGC CACATCGACA TTCAAGAATT
651 GCAGGAAAGG GCCGCGTCCG GCGCGGAATT GTCCACCACC GAAGCCCTGC
701 GCCTCGAATC CTTTGAAAAA GTCAACGCGC TGGGCATCGG CGCACAAGGC
751 TTGGGCGGAC TGACCACCGT GTTGGACGTG AAAATCCTCG ATTATCCGAC
801 CCACGCCGCC TCCAAACCGA TTGCCATGAT TCCGAATGCG GCCGCCACCC
851 GCCGCGACGT ATTTGAATTG GACGGCTCAG GCCCTGTCGA ACTCAGCCCG
901 CCGCGCGTCG AAGACTGGCC CGATTTGACT TACAGCCCCG ACAACGGCAA
951 ACGCGTCGAT GTCGACAAGC TGACCAAGA AGAAGTGGCA AGCTGGAAAA
1001 CGGCGGACGT ATTGCTGTTG AACGGCAAAA TCCTACCCGG CCGCGATGCC
1051 GCACACAACG GCCTCGTCGA TATGCTCAAC AAAGGCGAAG AATTGCCCGT
1101 CGATTTACAC AACCGCTCGA TTTACTACGT CGGCCCGCTC GATCCGGTCC
1151 CGGATGAAGT CGTCGGTCCG GCAGGTCCGA CCACAGCCAC CCGCATGGAC
1201 AAATTCACCC GCCAAATGCT CGAACAAACC GACCTCTTGG GCATGATCGG
1251 CCAATCCGAG CGCGGCGTGG CCACCTGCGA AGCCATCGCC GACAACAAAG
1301 CCGTGTACCT CATGGCAGTC GCGGGCGCGG CGTATCTCGT GGCAAAAGCC
1351 ATCAAATCTT CCAAAGTCTT GCGGTTCCCC GAATTGGGCA TGGAAGCCAT
1401 TTACGAATTT GAAGTCAAAG ACATGCCCGT AACCGTCGCC GTAGATAGCA
1451 AAGGCGAATC CATCCACGCC ACCGCCCCCG GCAATGGCA GCGGAAAATC
1501 GGCATCATCC CCGTCGAATC TTGA

```

This corresponds to the amino acid sequence <SEQ ID 220; ORF 050-1>:

m050-1.pep

```

1  MTVIKQEDFI QSIDAFQFI SYHPKDYID ALYKAWQKEE NPAAKDAMTQ
51  ILVNSRMCAE NNRPICQDTG IATVFLKVGM NVQWDADMSV EEMVNEGVR
101 AYTWEGNTRL ASVLADPAGK RQNTKDNTPA VIHMSIVPGG KVEVTCAAKG
151 GGSSENKSLA MLNPSDNIVD WVLKTIPTMG AGWCPPGILG IGIGGTPEKA
201 VLMAKESLMS HIDIQELQEK AASGAELSTT EALRLELF EK VNALGIGAQG
251 LGGLTTVL DV KILDYPTHA A SKPIAMIPNC AATRHVEFEL DGSGPVELTP
301 PRVEDWPDLT YSPDNGKRVD VDKLTKEEVA SWKTGDVLLL NGKILTRDA
351 AHKRLVMDLN KGEELPVDFE NRIYYVGPV DPVGDEVVGP AGPTTATRMD
401 KFTRQMLEQT DLLGMIGKSE RGVATCEAIA DNKAVYLM AV GGAAYLVAKA
451 IKSSKVLAF P ELGMEAIYEF EVKDMPTVTA VDSKGESIHA TAPRKWQAKI
501 GIIPVES*

```

m050-1/g050-1 98.2% identity in 507 aa overlap

	10	20	30	40	50	60
m050-1.pep	MTVIKQEDFIQSIDAFQFISYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE					
g050-1	MTVIKQEDFIQSIDAFQFISYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE					
	10	20	30	40	50	60
m050-1.pep	NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAITWEGNTRLASVLADPAGK					
g050-1	NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAITWEGNTRLASVLADPAGK					
	70	80	90	100	110	120
m050-1.pep	RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGSENKSLAMLNPSDNIVDWVLKTIPTMG					
g050-1	RQNTKDNTPAVIHMSIVPGGKVEVTCAAKGGSENKSLAMLNPSDNIVDWVLKTIPTMG					
	130	140	150	160	170	180
m050-1.pep	AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAAASGAELSTTEALRLELF EK					
g050-1	AGWCPPGILGIGIGGTPEKAVLMAKESLMSHIDIQELQEKAAASGAELSTTEALRLELF EK					
	190	200	210	220	230	240
m050-1.pep	VNALGIGAQGLGGLTTVL DVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP					
g050-1	VNALGIGAQGLGGLTTVL DVKILDYPTHAASKPIAMIPNCAATRHVEFELDGSGPVELTP					
	250	260	270	280	290	300

256

	310	320	330	340	350	360
m050-1.pep	PRVEDWPDLTYS	PDNGKRV	VDVDKLTKEE	VASWKTGDV	LLNGKILTGR	DAAHKRLVDMLN
g050-1	PRVEDXPDLTYS	PDNGKRV	VDVDKLTKEE	VASWKTGDV	LLNGKILTGR	DAAHKRLVNMLD
	310	320	330	340	350	360
	370	380	390	400	410	420
m050-1.pep	KGEELPVDF	TNRLIYYV	GPVDPVG	DEVVGPAG	PTTATRMD	KFTQMLQTDLLGMIGKSE
g050-1	KGEELPVDF	TNRLIYYV	GPVDPVG	DEVVGPAG	PTTATRMD	KFTQMLQTDLLGMIGKSE
	370	380	390	400	410	420
	430	440	450	460	470	480
m050-1.pep	RGVATCEAI	ADNKAVY	LMVGGAA	YLVAKAIK	SSKVLAF	PELGMEDIAIYEFV
g050-1	RGAATCEAI	ADNKAVY	LMVGGAA	YLVAKAIK	SSKVLAF	PELGMEDIAIYEFV
	430	440	450	460	470	480
	490	500				
m050-1.pep	VDSKGESI	HATAPRK	WQAKIGI	IPVESX		
g050-1	VDSKGESI	HATAPRK	WQAKIGI	IPVESX		
	490	500				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 221>:

a050-1.seq

1	ATGACCGTCA	TCAAACAGGA	AGACTTTATC	CAAAGCATTT	GCGATGCGCTT
51	CCAATTTCATC	AGCTACTACC	ATCCCAAAGA	CTACATCGAC	GCGCTTTATA
101	AGGCGTGGCA	GAAGGAAGAA	AACCCGCGCG	CCAAAGACGC	GATGACGCAG
151	ATTTTGGTCA	ACAGCCGCAT	GTGTGCCGAA	AACAACCGCC	CCATCTGCCA
201	AGATACCGGT	ATCGCGACCG	TGTTTTTGAA	AGTCGGTATG	GATGTGCAAT
251	GGGATGCAGA	CATGAGCGTC	GAAGAGATGG	TTAACGAAGG	CGTGCGCCGC
301	GCCTACACTT	GGGAAGGCAA	TACGCTGCGC	GCTTCCGTTC	TCGCCGACCC
351	CGCCGGCAAA	CGCCAAAATA	CCAAAGACAA	CACGCCCGCC	GTCATCCATA
401	TGAGCATCGT	GCCGGGCGAC	AAAGTCGAAG	TAACCTGCGC	GGCAAAAGGC
451	GGCGGTTCTG	AAAACAAATC	CAAACTCGCC	ATGCTCAACC	CTTCGACAA
501	CATCGTCGAT	TGGGTATTGA	AAACCATTC	GACCATGGGC	GCGGGCTGGT
551	GTCCCTCCCG	CATCTTGGGC	ATCGGCATCG	GCGGTACGCC	CGAAAAAGCC
601	GTGTTGATGG	CGAAAGAATC	CCTGATGAGC	CACATCGACA	TCCAAGAATT
651	GCAGGAAAAA	GCCGCGTCCG	GCGCGGAATT	GTCCACCACC	GAAGCCCTGC
701	GGCTCGAACT	CTTTGAAAAA	GTCAACGCGC	TAGGCATCGG	CGCGCAAGGC
751	TGGGCGGTG	TGACCACCGT	GTGGACGTG	AAAATCCTCG	ATTACCCGAC
801	CCACGCCGCC	TCCAAACCGA	TTGCCATGAT	TCCGAATGCG	GCCGCCACCC
851	GCCACGTGCA	ATTTGAATTG	GACGGCTCAG	GCCCTGTGCA	ACTCACGCCG
901	CGCGCGGTG	AAGACTGGCC	CGATTGACT	TACAGCCCGG	ACAACGGCAA
951	ACGCGTCGAT	GTCGACAAGC	TGACCAAAGA	AGAAGTGGA	AGCTGGAAAA
1001	CGGCGGACGT	ATTGCTGTTG	AACGGCAAAA	TCCTCACCGG	CCGCGATGCC
1051	GCACACAAAC	GCCTCGTCGA	TATGCTCGAC	AAAGGCGAAG	AATTGCCCGT
1101	CGATTTCACC	AACCGCCTGA	TTTACTACGT	CGGCCCGGTC	GATCCGGTCG
1151	GCGACGAAAT	CGTCGGCCCA	GCAGGTCCGA	CCACCGCCAC	CCGCATGGAC
1201	AAATTCACCC	GCCAAATGCT	CGAACAAACC	GACCTCTTGG	GCATGATCGG
1251	CAATCCGAG	CGCGGCGCGG	CCACCTGCGA	AGCCATCGCC	GACAACAAAG
1301	CCGTGTACCT	CATGGCAGTC	GCGGCGCGCG	CGTATCTCGT	GGCAAAAGCC
1351	ATCAAATCTT	CCAAAGTCTT	GGCGTTCCCC	GAATTGGGCA	TGGAAGCCAT
1401	TTACGAATTT	GAAGTCAAAG	ACATGCCCGT	AACCGTCGCC	GTAGACAGCA
1451	AAGGCGAATC	CATCCACGCC	ACCGCCCGCG	CCCAATGGCA	GGCGAAAAATC
1501	GGCATCATCC	CCGTCAAATC	TTGA		

This corresponds to the amino acid sequence <SEQ ID 222; ORF 050-1.a>:

a050-1.pep

1	MTVIKQEDFI	QSICDAFQFI	SYYPKDYID	ALYKAWQKEE	NPAAKDAMTQ
51	ILVNSRMCAE	NNRPICQDTG	IATVFLKVGM	DVQWDADMSV	EEMVNEGVR
101	AYTWEGNLT	ASVLADPAGK	RQNTKDNTPA	VIHMSIVPGD	KVEVTCAAKG
151	GGSENKSKLA	MLNPSDNIVD	WVLKTIPTMG	AGWCPPGILG	IGIGGTPEKA
201	VLMKESLMS	HIDIQELQEK	AASGAELSTT	EALRLELFEK	VNALGIGAQG
251	LGGLTTVL	LDV KILDYPT	HAA SKPIAMIP	NC AATRHVEFEL	DGSGPVELTP
301	PRVEDWPDLT	YSPDNGKRV	VDKLTKEEVA	SWKTGDVLLL	NGKILTGRDA
351	AKHRLVDMLD	KGEELPVDF	TNRLIYYVGP	VPVGVGDEIV	GP AGPTTATRMD
401	KFTQMLQET	DLLGMIGKSE	RGAATCEAIA	DNKAVYLMV	GGAAVLVAKA
451	IKSSKVLAFP	ELGMEAIYEF	EVKDMPTVA	VDSKGESIHA	TAPPQWQAKI
501	GIIPVKS*				

a050-1/m050-1 98.4% identity in 507 aa overlap

	10	20	30	40	50	60
a050-1.pep	MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE					
m050-1	MTVIKQEDFIQSICDAFQFISYYHPKDYIDALYKAWQKEENPAAKDAMTQILVNSRMCAE					
	10	20	30	40	50	60
	70	80	90	100	110	120
a050-1.pep	NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK					
m050-1	NNRPICQDTGIATVFLKVGMDVQWDADMSVEEMVNEGVRRAYTWEGNTLRASVLADPAGK					
	70	80	90	100	110	120
	130	140	150	160	170	180
a050-1.pep	RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG					
m050-1	RQNTKDNTPAVIHMSIVPGDKVEVTCAAKGGGSENKSKLAMLNPSDNIVDWVLKTIPTMG					
	130	140	150	160	170	180
	190	200	210	220	230	240
a050-1.pep	AGWCPPGILGIGGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLELFEK					
m050-1	AGWCPPGILGIGGGTPEKAVLMAKESLSHIDIQELQEKAASGAELSTTEALRLELFEK					
	190	200	210	220	230	240
	250	260	270	280	290	300
a050-1.pep	VNALGIGAQGLGGLTTVLVDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGGSGPVELTP					
m050-1	VNALGIGAQGLGGLTTVLVDVKILDYPHTAASKPIAMIPNCAATRHVEFELDGGSGPVELTP					
	250	260	270	280	290	300
	310	320	330	340	350	360
a050-1.pep	PRVEDWPDLTYSNDNGKRVVDVKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLD					
m050-1	PRVEDWPDLTYSNDNGKRVVDVKLTKEEVASWKTGDVLLNGKILTGRDAAHKRLVDMLN					
	310	320	330	340	350	360
	370	380	390	400	410	420
a050-1.pep	KGEELPVDFTNRLIYYVGPVDPVGDEIVGPAGPTTATRMDFTRQMLEQTDLLGMIGKSE					
m050-1	KGEELPVDFTNRLIYYVGPVDPVGDEVVGPAGPTTATRMDFTRQMLEQTDLLGMIGKSE					
	370	380	390	400	410	420
	430	440	450	460	470	480
a050-1.pep	RGAAATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVKDMPVTVA					
m050-1	RGVATCEAIADNKAVYLMVGGAAAYLVAKAIKSSKVLAFPELGMEAIYEFVKDMPVTVA					
	430	440	450	460	470	480
	490	500				
a050-1.pep	VDSKGESIHATAFPQWQAKIGIIPVKSX					
m050-1	VDSKGESIHATAPRKWQAKIGIIPVESX					
	490	500				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 223>:

g052.seq

```

1  ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTCAAAGG
51  CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
151 AAGGGGTTGG ACGGCGTTTC CAAAAACAGC AGTTTGGTGT TGGCTTTGAC
201 GGCGGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CTCACTCCGA
251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
351 CGCCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 224; ORF 052.ng>:

g052.pep

258

1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
 51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
 101 RLRLETTWSP ACRKVKNA*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 225>:

m052.seq
 1 ATGGCTTTGG TGGCGGAGGA AACGGAAATA TCCGCGCCGT GTTCAAAGG
 51 CTGCGAGCCG ACGGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
 101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCGCCC
 151 AAGGGGTTGG ACGGCGTTTC CAAAACAGC AGTTTGGTGT TGGCTTTGAC
 201 GCGCGCTTTC CATTCAATTA TATCAGTCGG CGACACGCGG CTCACTCCGA
 251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
 301 AGGCTGCGGC TGGAAACCAC ATGGTCGCCC GCCTGCAGGA AGGTGAAAAA
 351 CGCCGCCTGA

This corresponds to the amino acid sequence <SEQ ID 226; ORF 052>:

m052.pep
 1 MALVAEETEI SAPCFKGCEP TGDSRLLSTT KSAPMPCANS AKASKSATSP
 51 KGLDGVSKNS SLVLALTAAF HSFISVGDTR LTPMPNLVTM LLIKPTVVPN
 101 RLRLETTWSP ACRKVKNA*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 227>:

a052.seq
 1 ATGGCTTTGG TCGCGGAGGA AACGGAAATA TCCGCGCCGT GTTCAAAGG
 51 CTGAGAGCCG ACAGGCGACA GCAGGCTGTT GTCCACCACC AAGAGCGCGC
 101 CGATGCCGTG CGCCAATTCC GCCAAGGCTT CCAAGTCGGC CACTTCTCCC
 151 AAGGGATTGG ACGGCGTTTC CAAAACAGC AGTTTGGTGT TGGCTTTGAC
 201 GCGCGCTTTC CATTCGTTTA TATCAGTCGG CGACACGTA CTCACTCGA
 251 TGCCGAATTT GGTAACGATG TTATTGATAA AGCCGACGGT CGTGCCGAAC
 301 AGGCTGCGGC TGGAAATCAC ATGGTCGCCC GCCTGCAAAA AGGTGAAAAA
 351 CGCCGCCTGA

This corresponds to the amino acid sequence <SEQ ID 228; ORF 052.a>:

a052.pep
 1 MALVAEETEI SAPCFKG*EP TGDSRLLSTT KSAPMPCANS AKASKSATSP
 51 KGLDGVSKNS SLVLALTAAF HSFISVGD* LTSMPLVTM LLIKPTVVPN
 101 RLRLEITWSP ACKKVKNA*

m052/a052 95.8% identity over a 119 aa overlap

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
a052	MALVAEETEI SAPCFKGXEP TGDRLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m052.pep	SLVLALTAAFHSFISVGDTRLTPMPNLVTMLLIKPTVVPNRLRLLETTWSPACRKKVNAAX					
a052	SLVLALTAAFHSFISVGDXTLSMPNLVTMLLIKPTVVPNRLRLLEITWSPACKKVKNAAX					
	70	80	90	100	110	120

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 052 shows 95.8% identity over a 119 aa overlap with a predicted ORF (ORF 052.ng) from *N. gonorrhoeae*:

m052/g052

	10	20	30	40	50	60
m052.pep	MALVAEETEI SAPCFKGCEPTGDSRLLSTTKSAPMPCANS AKASKSATSPKGLDGVSKNS					

259

```

g052      MALVAEETEISAPCFKGCEPTGDSRLSTTKSAPMPCANSKASKSATSPKGLDGVSKNS
           10      20      30      40      50      60

           70      80      90      100     110     120
m052.pep  SLVLALTAAFHSFISVGDTWLTSMPLATMLLIKPTVVPNRLRLLEITWSPACKKVKNAAAX
           |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g052      SLVLALTAAFHSFISVGDTRLTPMPLVTMLLIKPTVVPNRLRLLETTWSPACRKVKNAAX
           70      80      90      100     110     120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 229>:

```

g073.seq
1  ATGTGTATGC CATACGCAAT AAGGGTTTCA GACGGCATCT GCCGCATTTT
51  TCCGCCGATG CCGTCTGAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
151 AGTCCGGGGC GGatacCGGC GCGGAGTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGGTT GTCCACGGAT TGGTGATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GAAAAGAGT TCCACGACTT TCCACGCGCG TGCTTGGTCG
301 GCGACTTCAA AACCGATGAC GATGCCGCGC CCGTTTTGCT GTTGC GGAT
351 AAGCTCCGCC TCGGATGGT CGGCAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 230; ORF 073.ng>:

```

g073.pep
1  MCMPYAIRVS DGICRIFPPM PSETRNQRAS ACFKSSIKSP TYSKPTDRRT
51  SPGRIPAASF SSGCILPCVV VHGLVMVERT SPRLAVREKS STTFHAAAWS.
101 ATSKPMTMPP PFCLRISSA CGWSGNPV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 231>:

```

m073.seq
1  ATGTGTATGC CATATAAGAT AAGGGTTTCA GACGGCATCT GCTGTCCAAT
51  GCCGTCTGAA ACACGCAATC AGCGTGCGAG TGCTGTTC AAATCGTCAA
101 TCAAATCGCC AACATATTCC AAACCGACCG ACAGGCGCAC CAATCCGGGG
151 CGGATGTTGG CCGCGAGTTT TTCTTCGGGC TGCATCCTGC CGTGCGTGGT
201 TGTCCACGGG TGGGTAATGG TCGAGCGCAC GTCACCGAGG TTGGCGGTGC
251 GGGAAAAGAG TTCCACGCCG TCCACAATT TCCACGCCGC TTCTTGATCG
301 GCAACTTCAA AGCCGATGAC GATGCCGCGC CCGTTTTGCT GTTGC GGAT
351 AAGCGCCGCC TGAGGATGGT CGGACAATCC GGTGTAG

```

This corresponds to the amino acid sequence <SEQ ID 232; ORF 073>:

```

m073.pep
1  MCMPYKIRVS DGICCPMPSE TRNQRASACF KSSIKSPTY S KPTDRRTNPG
51  RMLAASFSSG CILPCVVVHG WVMVERTSPR LAVREKSTP STTFHAASXS
101 ATSKPMTMPP PFCLRISAA XGWSNPNV*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 233>:

```

a073.seq
1  ACGTGTATGT CATATAAGAT AAGGGTTTCA GACGGCATTT GCGGTGTTTT
51  TCCGCCGATG CCGTCTGAA CACGCAATCA GCGCGCGAGT GCCTGTTTCA
101 AATCGTCAAT CAAATCGCCA ACATATTCCA AACCGACCGA CAGGCGCACC
151 AATCCGGGGC GGATGTTGGC GCGGAGTTT TCTTCGGGCT GCATCCTGCC
201 GTGCGTGGTT GTCCACGGAT GGGTAATGGT CGAGCGCACG TCGCCGAGGT
251 TGGCGGTACG GGAGAAAAGT TCGACGCCGT CCACGACTTT CCACGCGGCT
301 GCTTGGTCCG CCACTTCAA GCGGATGACG ATGCCGCCG CTTTTGCTG
351 TTTGCGGATA AGCTCCGCCT GAGGATGCTC GGGTAATCCG GTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 234; ORF 073.a>:

```

a073.pep
1  TCMSYKIRVS DGICGVFPPM PSEXRNQRAS ACFKSSIKSP TYSKPTDRRT
51  NPGRMLAASF SSGCILPCVV VHGWVMVERT SPRLAVREKS STPSTTFHAA
101 AWSATSKPMT MPPFFCLRI SSA*GWSGNP V*

```

m073/a073 92.3% identity over a 130 aa overlap

260

```

              10      20      30      40      50
m073.pep      MCMPIKIRVSDGICC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
              || ||||| ||||| |||||:||||| ||||| ||||| ||||| ||||| ||||| |||||
a073           TCMSYKIRVSDGICGVFPMPSEXRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
              10      20      30      40      50      60

              60      70      80      90      100     110
m073.pep      SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCCLRI
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a073           SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAAAWSATSKPMTMPPPFCCCLRI
              70      80      90      100     110     120

              120     129
m073.pep      SAAXGWSNPNVX
              |: ||||| |||||
a073           SSAXGWSGNPNVX
              130

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 073 shows 87.0% identity over a 131 aa overlap with a predicted ORF (ORF 073.ng) from *N. gonorrhoeae*:

m073/g073

```

              10      20      30      40      50
m073.pep      MCMPIKIRVSDGICC---PMPSETRNQRASACFKSSIKSPTYSKPTDRRTNPGRMLAASF
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g073           MCMPIAIRVSDGICRIFPMPSETRNQRASACFKSSIKSPTYSKPTDRRTSPGRIPAASF
              10      20      30      40      50      60

              60      70      80      90      100     110
m073.pep      SSGCILPCVVVHGWVMVERTSPRLAVREKSSTPSTTFHAASXSATSKPMTMPPPFCCCLRI
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g073           SSGCILPCVVVHGLVMVERTSPRLAVREKSST---TFHAAAWSATSKPMTMPPPFCCCLRI
              70      80      90      100     110

              120     129
m073.pep      SAAXGWSNPNVX
              |: ||||| |||||
g073           SSACGWSGNPNVX
              120

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 235>:

g075.seq

```

1  ATGCCGCCTT ACTTCATCAC CCTCTTAACG ATGGAATAA CAAAAAGCGC
51  GGCAGAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCCGTTTCGG GCAACGCTGC GTTTCCTGT
151 GCCGCCAAG CCAGCGGGGC GGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTCATGAT ATTTCCCTTT ACGAAATTT
251 TAAAAAATG TGTTCGCGG CTTGTGAAG GTTTAGAGA CCGCCTGCCG
301 GGCCTCTTAA ACTTAATCTT CTTTTTCGTA GAATCCGAAA ATTACAAATT
351 CCCGCCTAT CTCTCCAAT GCCGAGCTAA AAGCGTCTTC ATAGCTGTCA
401 TATTACCGG CTGA

```

This corresponds to the amino acid sequence <SEQ ID 236; ORF 075.ng>:

g075.pep

```

1  MPPYFILLT MENTKSAKT PTTIQPASIP SAFAASKAFF AVSGNAAFAC
51  AAKASGAAVT TASFAPYLQ VLINFMIFS TKFLKKCVCG LCEGFRDRLP
101 GLLNLIFFFV ESENYKFPAY LFQCRKSVF IAVIFTG*

```

m075/a075 98.5% identity over a 136 aa overlap

```

      10      20      30      40      50      60
m075.pep  MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFAAKARGAAVT
          |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a075      MPSYFITLLTMEKTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNVAFAAKARGAAVT
          10      20      30      40      50      60

      70      80      90     100     110     120
m075.pep  TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
          ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a075      TASFAPYLRQVLINFMIFSFKKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVADFFQT
          70      80      90     100     110     120

      130
m075.pep  CVNRFFEVEIIGIDX
          ||||| ||||| |||||
a075      CVNRFFEVEIIGIDX
          130

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 241>:

```

g080.seq
1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGGCATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAAGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA TACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGGTTGCA CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAA TGTCTTTTGG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACgaGA TGAACGCCT CCgGCTTTTT
601 ACcgAAGCGT GGCAGCATCT gttgcGTAAG AATAAAATC GGTATCCTA
651 TGTGGATATG Aggtataagg acggatttTC agtcccccat gctCCCGACG
701 GTTTACCCGA AAAAGAATcc gAAGAAtatt gggaaacaggt ttgggacata
751 ttacggcctg gcgtcggaac cggttcgacg caaatatcaa tcagttatAA
801 GGGCAGacga acaatggaac AGcagtaa

```

This corresponds to the amino acid sequence <SEQ ID 242; ORF 080.ng>:

```

g080.pep
1  MWDNAEAMER LTRWLLVMMa MLLAASGLVW FYNSNHLVPK QVSLKGNLVY
51  SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDIVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWNVVL DNGITVRLGR ENEMKRLRLF
201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVPH APDGLPEKES EEYWEQVWDI
251 LRPGVGNST QISISYKGRR TMEQQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 243>:

```

m080.seq
1  ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
51  CATGATGGCG ATGCTGCTTG CTGCGTCCGG GCTGGTTTGG TTTTACAATT
101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTGGTTTAT
151 TCCGATAAGA AGACATTGGG CAGTTTGGCG AAAGAATACA TCCATGGGAA
201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
301 GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT
351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGCTTGGAC AGACCCGGAA
401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCTTTTGG GACAACGGCA
551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAACGCCT CCgGCTTTTT

```

601 ACCGAAGCGT GGCAGCATCT GTTGCCTAAA AATAAAAATC GGTATCCTA
 651 TGTGGATATG AGGTATAAGG ACGGATTTTC AGTCCGCTAT GCTTCCGACG
 701 GTTTACCCGA AAAAGAATCC GAAGAATAG

This corresponds to the amino acid sequence <SEQ ID 2441; ORF 080>:

m080.pep
 1 MWDNAEAMER LTRWLLVMM MLLAASGLVW FYNSNHLVPK QVSLKGNLVY
 51 SDKKTGSLA KEYIHGNILR TDINGAQEAY RRYPWIASVM VRRRFPDVE
 101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
 151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
 201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY ASDGLPEKES EE*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 080 shows 97.9% identity over a 242 aa overlap with a predicted ORF (ORF 080.ng) from *N. gonorrhoeae*:

m080/g080

m080.pep	10	20	30	40	50	60
	MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLVPKQVSLKGNLVYSDKKTGSLA					
080	MWDNAEAMERLTRWLLVMMAMLLAASGLVWFYNSNHLVPKQVSLKGNLVYSDKKALGSLA					
	10	20	30	40	50	60
m080.pep	70	80	90	100	110	120
	KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDVEVVLTERKPVARWGDHALVDG					
080	KEYIHGNILRTDINGAQEAYRRYPWIASVMVRRRFPDVEVVLTERKPVARWGDHALVDG					
	70	80	90	100	110	120
m080.pep	130	140	150	160	170	180
	EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWIVVL					
080	EGNVFEARLDRPGMPVFRGAEGTSAEMLRRYDEFSTVLAKQGLGIKEMTYTARSAWNVVL					
	130	140	150	160	170	180
m080.pep	190	200	210	220	230	240
	DNGITVRLGRENMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVRYASDGLPEKES					
080	DNGITVRLGRENMKRLRLFTEAWQHLLRKNKNRLSYVDMRYKDGFSVPHAPDGLPEKES					
	190	200	210	220	230	240
m080.pep	EEX					
080	EEYWEQVVDILRPGVGNSTQISISYKGRRTMEQQX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 245>:

a080.seq
 1 ATGTGGGATA ATGCCGAAGC GATGGAACGG CTGACGCGCT GGCTGCTTGT
 51 CATGATGCGC ATGCTGCTTG CTGCGTCCGG GCTGTTTGG TTTTACAATT
 101 CGAATCATCT GCCCGTCAAG CAGGTGTCGC TGAAGGGCAA CCTAGTTTAT
 151 TCCGATAAGA AAGCATTGGG CAGTTTGCGC AAAGAATACA TCCATGGGAA
 201 TATTTTGAGG ACGGACATCA ATGGCGCACA GGAGGCCTAC CGCCGGTATC
 251 CGTGGATTGC GTCGGTCATG GTGCGCCGCC GTTTTCCCGA CACGGTTGAG
 301 GTCGTCCTGA CCGAGCGCAA GCCGGTCGCG CGTTGGGGCG ACCATGCCTT
 351 GGTGGACGGC GAAGGCAATG TTTTGAAGC CCGTTTGGAC AGACCCGGAA
 401 TGCCGGTATT CAGAGGCGCG GAAGGAACGT CTGCCGAAAT GCTCCGCCGT
 451 TATGACGAAT TTTCGACTGT TTTGGCAAAA CAGGGTTTGG GCATCAAAGA
 501 GATGACCTAT ACGGCACGTT CGGCGTGGAT TGTCGTTTTG GACAACGGCA

264

```

551 TCACCGTCAG GCTCGGACGG GAAAACGAGA TGAAACGCCT CCGGCTTTTT
601 ACCGAAGCGT GGCAACATCT GTTGCCTAAA AATAAAATC GGTTATCCTA
651 TGTGGATATG AGGTATAAGG ACGGATTTC AGTCCGCTAT GCTCCCGACG
701 GTTACCCGA AAAAGAATCC GAAGAATAG

```

This corresponds to the amino acid sequence <SEQ ID 246; ORF 080.a>:

a080.pep

```

1 MWDNAEAMER LTRWLLVMM MLLAASGLVW FYNHLPVK QVSLKGNLVY
51 SDKKALGSLA KEYIHGNILR TDINGAQEAY RRYPIASVM VRRRFPDVE
101 VVLTERKPVA RWGDHALVDG EGNVFEARLD RPGMPVFRGA EGTSAEMLRR
151 YDEFSTVLAK QGLGIKEMTY TARSAWIVVL DNGITVRLGR ENEMKRLRLF
201 TEAWQHLLRK NKNRLSYVDM RYKDGFSVRY APDGLPEKES EE*

```

m080/a080 99.2% identity over a 242 aa overlap

	10	20	30	40	50	60
m080.pep	MWDNAEAMERL	TRWLLVMMAM	LLAASGLVWF	FYNHLPVKQV	SLKGNLVYS	DKKTLGSLA
a080	MWDNAEAMERL	TRWLLVMMAM	LLAASGLVWF	FYNHLPVKQV	SLKGNLVYS	DKKALGSLA
	10	20	30	40	50	60
	70	80	90	100	110	120
m080.pep	KEYIHGNILR	TDINGAQEAY	RRYPWIASVM	VRRRFPDVE	VVLTERKPVA	RWGDHALVDG
a080	KEYIHGNILR	TDINGAQEAY	RRYPWIASVM	VRRRFPDVE	VVLTERKPVA	RWGDHALVDG
	70	80	90	100	110	120
	130	140	150	160	170	180
m080.pep	EGNVFEARLD	RPGMPVFRGA	EGTSAEMLRR	YDEFSTVLAK	QGLGIKEMTY	TARSAWIVVL
a080	EGNVFEARLD	RPGMPVFRGA	EGTSAEMLRR	YDEFSTVLAK	QGLGIKEMTY	TARSAWIVVL
	130	140	150	160	170	180
	190	200	210	220	230	240
m080.pep	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLRK	NKNRLSYVDM	RYKDGFSVRY	APDGLPEKES
a080	DNGITVRLGR	ENEMKRLRLF	TEAWQHLLRK	NKNRLSYVDM	RYKDGFSVRY	APDGLPEKES
	190	200	210	220	230	240
m080.pep	EEX					
a080	EE					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 247>:

g081.seq

```

1 ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51 GCCGTCTGAA AACAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGATA
101 TTCGGGAAGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGACGCG
151 CATGATTTTG TTGGAGGCGT ATTGTCTGCG GGCGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGCGCGG CTTTGGGCGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC GTTGCAAACG TTGGCGAAGG CGTGGCGCGA TAATGTGAAC
301 CCGTTTGTGT TCGGCATTAC CGGTCGGGCG GCAAGACGA CCGTGAAGGA
351 GATGCTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATGAT GCCGTTTCGG
401 CGACGGCAGG CAACTTCAAC AACCACAtcg gaTTGCCGCT GACTTTATTG
451 AAATtaaAcg aAAAAACCCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTTGGcgaa ctggcggtTt taacgcaaat CGCCAAACCC GATGCCGCTT
551 TGGtcaACAA CGCCTGCGC GCCCATGTCG GATGCGGTTt cgacggagtg
601 GGCATATTG CCAAAGcgaa aagcGAGATT TatgcagGct tATGTTCAGA
651 CGGCATGGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGTT TAATTTGAAT ACGTGCACTT TCGGCGTCGA TAGCGCGAT
751 GTCCGCGCGG AAAATATCGT GCTGAAACCT TTGTCGTGCG AATTTGATT

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 237>:

```
m075.seq
1  ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAATA CAAAAGCGC
51  GGCGAAAATG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCGGTATCGG GCAACGTTGC ATTTGCATGT
151 GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTC AAAAAGTGTT
251 TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301 TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351 CTTCCAAACC TGCCTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401 TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 238; ORF 075>:

```
m075.pep
1  MPSYFITLLT MENTKSAAKM PTTIQPASIP SAFAASKAFF AVSGNVAFAC
51  AAKARGAAVT TASFAPYLRL VLINFMIFSF KKCLAVMDGA FFRRPPNIRK
101 SVFQKSEYDK FVLVADFFQT CVNRFFEVEE IIGIGD*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 075 shows 65.7% identity over a 137 aa overlap with a predicted ORF (ORF 075.ng) from *N. gonorrhoeae*:

```
m075/g075

          10      20      30      40      50      60
m075.pep  MPSYFITLLTMENTKSAAKMPTTIQPASIPSAFAASKAFFAVSGNVAFACA AKARGAAVT
          |||||||||||||||||||||||||||||||||||||||||||||
g075       MPPYFITLLTMENTKSAAKTPTTIQPASIPSAFAASKAFFAVSGNAAFACA AKASGAAVT
          10      20      30      40      50      60

          70      80      90      100     110
m075.pep  TASFAPYLRLQVLINFMIFSF---KKCLAVMDGAFFRRPPNIRKSVFQKSEYDKFVLVAD
          ||||||||||||||||||||| |||: : :| | : : :| | : : :|
g075       TASFAPYLRLQVLINFMIFSFTKFLKKVCVGLCEGFRDRLPGLLNLIFFVESENYKFPAY
          70      80      90      100     110     120

          120     130
m075.pep  FFQTCVNRFFEVEEIIIGIDX
          :|| : : | : : |
g075       LFQCRKSVFIAVIFTGX
          130
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 239>:

```
a075.seq
1  ATGCCGTCTT ACTTCATCAC TCTCTTAACG ATGGAAAAGA CAAAAGCGC
51  GGCGAAAACG CCCACTACAA TCCAACCGGC TTCCATACCG TCCGCTTTTG
101 CGGCTTCCAA AGCGTTTTTT GCTGTATCGG GCAACGTTGC ATTTGCATGT
151 GCGGCCAAAG CCAGGGGAGC AGCTGTTACA ACAGCCAGTT TTGCGCCGTA
201 TTTACGGCAG GTGTTAATAA ATTTTCATGAT ATTTTCCTTC AAAAAGTGTT
251 TGGCGGTAAT GGATGGAGCG TTTTTCAGAC GACCGCCGAA CATCCGAAAA
301 TCAGTCTTTC AAAAATCCGA ATACGACAAA TTCGTATTGG TTGCCGATTT
351 CTTCCAAACC TGCCTTAATC GCTTCTTCGA AGTCGTAGAA ATAATCGGCA
401 TTGGTGATTA A
```

This corresponds to the amino acid sequence <SEQ ID 240; ORF 075.a>:

```
a075.pep
1  MPSYFITLLT MEKTKSAAKT PTTIQPASIP SAFAASKAFF AVSGNVAFAC
51  AAKARGAAVT TASFAPYLRL VLINFMIFSF KKCLAVMDGA FFRRPPNIRK
101 SVFQKSEYDK FVLVADFFQT CVNRFFEVEE IIGIGD*
```

265

```

801 GGTGTGCGGC GACGAGCGCA CTGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG CAAGGCTTCA GCAACATCAA
951 AGGCCGTCTG AACGTCAAAG CCGGCATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAATCCC GACAGTATGA AAGCCGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG CATTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAGGAc gaAGCCGCCG CCATGCACGC CGAagtcgGC GCGTACGCCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 GCGGcggaAA AATTTGGCGC GGACGGTTTG TGGTTCGCCG CCAAAGACCC
1251 GTTGATTCAA GTGTGAGCC ACGATTGGCC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTT GCGCTTTATG CAGAtggaAG AAGTGGTCGA GGCATTGGAG
1351 GATAAGTga

```

This corresponds to the amino acid sequence <SEQ ID 248; ORF 081.ng>:

```

g081.pep
  1 MKPLDLNFIC QALKLPMPSE NKPVSRIVTD SRDIREGDVF FALAGGRFDA
 51 HDFVGGVLSA GAAAVVVSRE DCAALGGALK VDDTLAALQT LAKAWRDNVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVSATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTQIAKP DAALVNNALR AHVGCDFDGV
201 GDIKAKAKSEI YAGLCSDGMA LIPQEDANMA VFKTATFNLN TCTFGVDSGD
251 VRAENIVLKP LSCEFDLVCG DERTAVVLPV PGRHNVHNAA AAAALALAAg
301 LSLNDVAEGL QGFSNIKGR L NVKAGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQGIEA AYFVGDNVSE
401 AAEKFGADGL WFAAKDPLIQ VLSHDLPERA TVLVKGSRFM QMEEVVEALE
451 DK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 249>:

```

m081.seq
  1 ATGAAACCAC TGGACCTAAA TTTTCATCTGC CAAGCCCTCA AGCTTCCGAT
 51 GCCGTCTGAA AGCAAACCCG TGTGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGCGAGCG GTTTGACGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCT GGTGCGGCGG CGGTTGTGGT
201 TTCGCGCGAA GATTGTGCTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTGCCGC ATTGCAAAACG CTGGCAAAGG CGTGGCGTGA AAATGTGAAT
301 CCGTTTGTGT TCGGCATTAC CGGTTGCGGC GGCAAGACGA CGGTGAAGGA
351 AATGCTGGCT GCGTATTGCG GCCGCCGTTT CCGCGATGAT GCCGTGTTGG
401 CGACGGCAGG CAACTTCAAC AACCATATCG GATTGCCGCT GACTTTGTTG
451 AAGTTAAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GCATGAACCA
501 TTTCCGCGAA CTGGCGGTTT TAACGCmAT CGCCAAACCA AATGCCGCAT
551 TGGTCAACAA CGCCATGCGC GCCCATGTGCG GCTGCGGTTT CGACGGAGTG
601 GGCGATATTG CCAAAGCGAA AAGCGAGATT TACCAAGGTT TATGTTTCAGA
651 CGGCATTGCA CTGATTCCTC AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTTGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTTCACGCGG AAAATATTGT GCTGAAACCG TTGTCGTGCG AATTTGATTT
801 GGTGTGCGGC GATGAGCGCG CCGCCGTGGT GCTGCCTGTT CCCGGCCGCC
851 ACAATGTCCA CAACGCCGCC GCTGCCGCCG CGCTGGCTTT GGCTGCCGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGTCTG AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCCT GACAGCATGA AAGCTGCGAT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGCGAACTG GCGGAGGACG AAGCCGCCGC TATGCACGCC GAAGTCGGCG
1151 CGTATGCCCC CGACCAAGGC ATCGAAGCGG CTATTTTGTG CGGCGACAAC
1201 AGCGTCGAAG CGGCGGAAAA ATTTGGCGCG GACGGTTTGT GGTTCGCCGC
1251 CAAAGACCCG TTGATTCAAG TGTGCGCCA CGATTTGCCC GAACGCGCCA
1301 CCGTGTGGT GAAAGGTTTC CGCTTTATGC AGATGGAAGA AGTGGTCGAG
1351 GCATTGGAGG ATAAGTGA

```

This corresponds to the amino acid sequence <SEQ ID 250; ORF 081>:

```

m081.pep
  1 MKPLDLNFIC QALKLPMPSE SKPVSRIVTD SRDIRAGDVF FALAGERFDA
 51 HDFVEDVLAA GAAAVVVSRE DCAAMDGALK VDDTLAALQT LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDD AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFGE LAVLTXIAKP NAALVNNAMR AHVGCDFDGV

```

```

201  GDIKAKSEI YQGLCSDGIA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251  VHAENIVLKP LSCEFDLVCG DERAADVLPV PGRHNVHNAA AAAALALAAG
301  LSLNDVAEGL KGFSNIKGR LNVKSGIKGAT LIDDTYNANP DSMKAAIDVL
351  ARMPAPRIFV MGD MGELGEL GEDEAAAMHA EVGAYARDQG IEAAYFVGDN
401  SVEAAEKFGA DGLWFAAKDP LIQVLRHDLR ERATV LVKGS RFMQMEEVVE
451  ALEDK*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 081 shows 94.1% identity over a 455 aa overlap with a predicted ORF (ORF 081.ng) from *N. gonorrhoeae*:

m081/g081

m081.pep	10	20	30	40	50	60	
	MKPLDLNFICQALKLPMPSESKPVSRI	VTDSRDIRAGDVFFALAGERF	DAHDFVEDVLAA				
g081	10	20	30	40	50	60	
	MKPLDLNFICQALKLPMPSENKPVSRI	VTDSRDI	REGDVFFALAGGRF	DAHDFVGGVLSA			
m081.pep	70	80	90	100	110	120	
	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA						
g081	70	80	90	100	110	120	
	GAAAVVVSREDCAALGGALKVDDTLAALQTLAKAWRDNVNPFVFGITGSGGKTTVKEMLA						
m081.pep	130	140	150	160	170	180	
	AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNKHYAVIEMGMNHFGE	LA	VLTXIAKP				
g081	130	140	150	160	170	180	
	AVLRRRFGDDAVSATAGNFNNHIGLPLTLLKLNKHYAVIEMGMNHFGE	LA	VLTXIAKP				
m081.pep	190	200	210	220	230	240	
	NAALVNNAMRAHVGC	FDGVDIAKAKSEIYQGLCSDGIALIPQEDANMA	VF	KTATLNLN			
g081	190	200	210	220	230	240	
	DAALVNNALRAHVGC	FDGVDIAKAKSEIYAGLCSDGMALIPQEDANMA	VF	KTATLNLN			
m081.pep	250	260	270	280	290	300	
	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVLPVPGRHNVHNAAAAA	LALAAG					
g081	250	260	270	280	290	300	
	TCTFGVDSGDVRAENIVLKPLSCEFDLVCGDERTAVVLPVPGRHNVHNAAAAA	LALAAG					
m081.pep	310	320	330	340	350	360	
	LSLNDVAEGLKGFSNIKGR LNVKSGIKGATLIDDTYNANPDSMKAAIDVL	ARM	PAPRIFV				
g081	310	320	330	340	350	360	
	LSLNDVAEGLQGFSNIKGR LNVKAGIKGATLIDDTYNANPDSMKAAVDVL	ARM	PAPRIFV				
m081.pep	370	380	390	400	410	420	
	MGDMGELGELGEDEAAAMHAEV	GAYARDQGEAAYFVG	DNSVEAAEKFGADGLWFAAKDP				
g081	370	380	390	400	410	420	
	MGDMGELGE---DEAAAMHAEV	GAYARDQGEAAYFVG	DNSVEAAEKFGADGLWFAAKDP				
m081.pep	430	440	450				
	LIQVLRHDLPERATV LVKGS RFMQMEEVVEALEDKX						
g081	420	430	440	450			
	LIQVLSHDLPERATV LVKGS RFMQMEEVVEALEDKX						

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 251>:

```
a081.seq
1  ATGAAACCAC TGGACCTAAA TTTCATCTGC CAAGCCCTCA AGCTTCCGAT
51  GCCGTCTGAA AGCAAACCCG TGTCGCGCAT CGTAACCGAC AGCCGCGACA
101 TCCGCGCGGG CGATGTGTTT TTCGCATTGG CGGGCGGGCG GTTTGATGCG
151 CATGATTTTG TTGAAGACGT ATTGGCTGCG GGTGCGGCGG CGGTGTGGT
201 TTCGCGCGAA GATTGCGTTG CAATGGATGG CGCGTTGAAA GTCGATGACA
251 CGCTTACCGC GTTGCAAATG TTGGCGAAGG CGTGGCGCGA GAATGTGAAC
301 CCGTTTGTGT TCGGTATTAC CGGCTCGGGC GGCAAGACGA CGGTGAAGGA
351 AATGTTGGCT GCGGTATTGC GCCGCCGTTT CGGCGATAAT GCCGTTTGG
401 CGACGCGCAG CAACTTCAAC AACCACATCG GATTGCCGTT GACTTTGTTG
451 AAATTAACG AAAAACACCG CTATGCCGTG ATTGAAATGG GTATGAACCA
501 TTTTGGCGAA CTGGCGGTTT TGACACAAAT CGCCAAACCC GATGCCGCAT
551 TGGTCAACAA CGCATGCGC GCCCATGTCG GCTGCGGTTT CAGCGGAGTG
601 GCGGATATTG CCAAAGCGAA AAGCGAGATT TATCAAGGCT TATGTTTCTG
651 CCGCATGGCG CTGATTCTCT AAGAAGATGC CAATATGGCT GTCTTCAAAA
701 CGGCAACGCT TAATTGGAAT ACGCGCACTT TCGGCATCGA TAGCGGCGAT
751 GTCCACGCGG AAAATATCGT GCTGAAACCG TTGTCGTGCG AATTGATTT
801 GGTGTGCGGC AACGAGTGGC CAGCCGTGGT TCTGCCCGTT CCCGCCGCC
851 ACAATGTCCA CAACGCCGCC GCCGCCGCCG CGCTGTCTTT GGCTGCAGGT
901 TTGAGTTTGA ACGATGTGGC GGAAGGTTTG AAAGGCTTCA GCAATATCAA
951 AGGCCGCTCT AACGTCAAAT CCGGAATCAA GGGCGCAACC CTGATTGACG
1001 ATACTTATAA TGCGAACCC TACAGCATGA AAGCTGCGGT TGACGTGTTG
1051 GCGCGTATGC CTGCGCCGCG TATTTTCGTG ATGGGCGATA TGGGCGAACT
1101 GGGTGAGGAC GAAGCCGCCG CCATGCACGC CGAAGTCGGC GCGTACGCC
1151 GCGACCAAGG CATCGAAGCG GCTTATTTTG TCGGCGACAA CAGCGTCGAA
1201 CCGGCGGAAA AATTTGGCGC GGACGGTTTG TGGTTCGCGC CCAAAGACCC
1251 GTTGATTCAA GTGTGCGGCC ACGATTTGCC CGAACGCGCC ACCGTGTTGG
1301 TGAAAGGTTT CCGCTTTATG CAGATGGAAG AAGTGGTCGA GGCATTGGAG
1351 GATAAGTGA
```

This corresponds to the amino acid sequence <SEQ ID 252; ORF 081.a>:

```
a081.pep
1  MKPLDLNFIC QALKLMPSE SKPVSRIVTD SRDIRAGDVF FALAGGRFDA
51  HDFVEDVLAA GAAAVVVSRE DCVAMDGALK VDDTLTALQM LAKAWRENVN
101 PFVFGITGSG GKTTVKEMLA AVLRRRFGDN AVLATAGNFN NHIGLPLTLL
151 KLNEKHRYAV IEMGMNHFG E LAVLTQIAKP DAALVNNAMR AHVCGFDGV
201 GDIKAKAKSEI YQGLCSDGMA LIPQEDANMA VFKTATLNLN TRTFGIDSGD
251 VHAENIVLKP LSCEFDLVCG NECAAVVLPV PGRHNVHNAA AAAALSLAAG
301 LSLNDVAEGL KGFSNIKGR L NVKSGIKGAT LIDDTYNANP DSMKAAVDVL
351 ARMPAPRIFV MGDMGELGED EAAAMHAEVG AYARDQGIEA AYFVGDNVSE
401 AAEKFGADGL WFAAKDPLIQ VLRHDLPERA TVLVKGSREM QMEEVVEALE
451 DK*
```

m081/a081 96.7% identity over a 455 aa overlap

	10	20	30	40	50	60
m081.pep	MKPLDLNFICQALKLMPSESKPVSRIVTDSDRDIRAGDVFFALAGERFDAHDFVEDVLAA					
a081	MKPLDLNFICQALKLMPSESKPVSRIVTDSDRDIRAGDVFFALAGGRFDAHDFVEDVLAA					
	10	20	30	40	50	60
m081.pep	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
a081	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
	70	80	90	100	110	120
m081.pep	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
a081	GAAAVVVSREDCAAMDGALKVDDTLAALQTLAKAWRENVNPFVFGITGSGGKTTVKEMLA					
	130	140	150	160	170	180
m081.pep	AVLRRRFGDDAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTQIAKP					
a081	AVLRRRFGDNAVLATAGNFNNHIGLPLTLLKLNEKHRYAVIEMGMNHFGELAVLTQIAKP					

268

	130	140	150	160	170	180
m081.pep	190	200	210	220	230	240
	NAALVNNAMRAHVCGCFDVGVDIAKAKSEIYQGLCSGIALIPQEDANMAVFKTATLNLN					
a081	DAALVNNAMRAHVCGCFDVGVDIAKAKSEIYQGLCSGDMALIPQEDANMAVFKTATLNLN					
	190	200	210	220	230	240
m081.pep	250	260	270	280	290	300
	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGDERAAVVLVPVGRHNVHNAAAAAALALAAG					
a081	TRTFGIDSGDVHAENIVLKPLSCEFDLVCGNECAAVVLVPVGRHNVHNAAAAAALSALAAG					
	250	260	270	280	290	300
m081.pep	310	320	330	340	350	360
	LSLNDVAEGLKGFSNIKGRNLNVKSGIKGATLIDDTYNANPDSMKAAIDVLARMPAPRIFV					
a081	LSLNDVAEGLKGFSNIKGRNLNVKSGIKGATLIDDTYNANPDSMKAAVDVLARMPAPRIFV					
	310	320	330	340	350	360
m081.pep	370	380	390	400	410	420
	MGDMGELGELGEDEAAAMHAEVGAYARDQGIEAAYFVGDNVSEAAEKFGADGLWFAAKDP					
a081	MGDMGELGE---DEAAAMHAEVGAYARDQGIEAAYFVGDNVSEAAEKFGADGLWFAAKDP					
	370	380	390	400	410	
m081.pep	430	440	450			
	LIQVLRHDLPERATVVLVKGSRFMQMEEVVEALEDKX					
a081	LIQVLRHDLPERATVVLVKGSRFMQMEEVVEALEDKX					
	420	430	440	450		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 253>:

g082.seq

```

1  aTGTGGTTGT TGAAGTTGCC TGCCGTCGCC GAAACGGCAT CATCGCCGAA
51  ACGGCGGCGC AATACCGCAG CCAGCATCTC CTTACCGTC GTCTTGCCGC
101 CCGAACCGGT AATGCCGAAC ACAACGGGT TCACATTATC GCGCCACGCC
151 TTCGCCAACG TTTGCAACGC GGCAAGCGTG TCATCGACTT TCAACGCGCC
201 GCCCAAAGCC GCGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCGCCCG
251 CAGACAATAC GCCTCCAACA AAATCATGCG CGTCAAACCG CCCGCCCGCC
301 AATGCGAAAA ACACATCGCC TTCCCGAATA TCGCGGCTGT CGGTTACGAT
351 GCGCGACACG GTTTGTGTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401 AGATGAAATT TAGGTCCAGT GGTTCATAT TTGCTTTCGT TAATATTCGG
451 GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501 GGTA AACAC AGATTATTTT CCCATTCTCA TTCGGCATT TTTCTGTACG
551 TATCATTTT TAGACGTATT TTTAGCCGAT TTGCCTTTTC CCGCATACCA
601 CGGCGCGGGG TCGTCGGACT GTCTGTCGAT AAAGCAAGG TTATTGCCTT
651 CGCCCGGCAC ATCGGGGACA TTCCCCCAA AATCATAGCC GTCATCGGGC
701 AACTCGTCGG TTTGATACC CGTCCAACG CCGAATCCG GTAA

```

This corresponds to the amino acid sequence <SEQ ID 254; ORF 082.ng>:

g082.pep

```

1  MWLLKLPAVA ETASSPKRRR NTAASISFTV VLPPEPVMPN TNGFTLSRHA
51  FANVCNAASV SSTFNAPPKA AQSSRETTTA AAPADNTPPT KSCASNRPFA
101 NAKNTSPSRI SRLSVTMRDT GLFSDGIGSL RAWQMKFRSS GFIFAFVNIR
151 AADTSVAADF FIACFAVVKH RLFSSHSAF FLYVSFFRRI FSRFAFSRIP
201 RRGVVGLSVD KGKVIARH IGDIPKIIA VIGQLVGFTD RPTAESA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 255>:

m082.seq

```

1  ATGnnGTTGT TGAAGTTGCC TGCCGTCGCC AACACGGCAT CATCGCCGAA
51  ACGGCGGCGC AATACCGCAG CCAGCATTTT CTTACCGTC GTCTTGCCGC
101 CCGAACCGGT AATGCCGAAC ACAACGGAT TCACATTTTC ACGCCACGCC
151 TTTGCCAGCG TTTGCAATGC GGCAAGCGTG TCATCGACTT TCAACGCGCC

```

269

```

201 ATCCATTGCA GCACAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCAG
251 CAGCCAATAC GTCTTCAACA AAATCATGCG CGTCAAACCG CTCGCCCCGC
301 AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTTACGAT
351 GCGCGACACG GGTTCGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGGC
401 AGATGAAATT TAGGTCCAGT GGTTCATAT TTACTTTCGT TAATATTCGG
451 GCGCGGACAC CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501 GGTAAACAC AGATTATTTT CCCATTCTCA TTCGGSATTT TTTCTGTACG
551 TATCATTTTT TAGACGTATT TTAGTCGAT TTGCCTTTTC CCGCATACCA
601 CGGCGCGGGG TCGTCGGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
651 CGCCTGCAC ATCGGGAACA TTCCCCAAA AATCATAGCC GTCATCGGGC
701 AACTCGTCGG TTCGATACC CGTCCAACG CCGAATCCGC GTAA

```

This corresponds to the amino acid sequence <SEQ ID 256; ORF 082>:

```

m082.pep
  1 MXLLKLPAVA NTASSPKRRR NTAASISFTV VLPPEPVM PN TNGFTFSRHA
 51 FASVCNAASV SSTFNAPSIA AQSSRETTA AAPAANTSST KSCASNRS PA
101 NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151 AADTSVAADF FIACFAVVKH RLFSHSHSXF FLYVSFFRRI FSRFAFSRIP
201 RRGVVGQSDV KGKVI AFALH IGNIPPKIIA VIGQLVGFDT RPTAESAX

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 082 shows 92.7% identity over a 247 aa overlap with a predicted ORF (ORF 082.ng) from *N. gonorrhoeae*:

```

m082/g082
      10      20      30      40      50      60
m082.pep MXLLKLPAVANTASSPKRRRNTAASISFTVLPPEPVMPTNGFTFSRHAFAFASVCNAASV
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g082      MWLLKLPAVAETASSPKRRRNTAASISFTVLPPEPVMPTNGFTLSRHAFAFANVCNAASV
          10      20      30      40      50      60

      70      80      90     100     110     120
m082.pep SSTFNAPSIAAQSSRETTAAAPANTSSTKSCASNRS PANAKNTSPARMSRLSVTMRDT
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g082      SSTFNAPPKAAQSSRETTAAAPADNTPPTKSCASNRP PANAKNTSPSRISRLSVTMRDT
          70      80      90     100     110     120

      130     140     150     160     170     180
m082.pep GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF
          ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g082      GLFSDGIGSLRAWQMKFRSSGFIFAFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF
          130     140     150     160     170     180

      190     200     210     220     230     240
m082.pep FLYVSFFRRIFSRFAFSRIPRRGVVGQSDVKGKVI AFALHIGNIPPKIIAVIGQLVGFDT
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g082      FLYVSFFRRIFSRFAFSRIPRRGVVGLSVDKGKVI AFARHIGDIPPKIIAVIGQLVGFDT
          190     200     210     220     230     240

m082.pep RPTAESAX
          |||||
g082      RPTAESAX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 257>:

```

a082.seq
  1 ATGTGGTTGT TGAAGTTGCC TGCCGTCGCC AAAACGGCAT TATCGCCGAA
 51 ACGGCGGCGC AATACGCAG CCAACATTTC CTTCACCGTC GTCTTGCCGC

```

270

```

101 CCGAGCCGGT AATACCGAAC ACAAACGGGT TCACATTCTC GCGCCACGCC
151 TTCGCCAACA TTTGCAACGC GGTAAGCGTG TCATCGACTT TCAACGCGCC
201 ATCCATTGCA ACGCAATCTT CGCGCGAAAC CACAACCGCC GCCGCACCCG
251 CAGCCAATAC GTCTTCAACA AAATCATGCG CATCAAACCG CCCGCCCGCC
301 AATGCGAAAA ACACATCGCC CGCGCGGATG TCGCGGCTGT CGGTACGAT
351 GCGCGACACG GGTTTGCTTT CAGACGGCAT CGGAAGCTTG AGGGCTTGCG
401 AGATGAAATT TAGGTCCAGT GGTTCATAT TTACTTTCGT TAATATTCGG
451 GCGGCGGACA CATCGGTAGC GGCTGATTTT TTTATCGCCT GTTTTGCTGT
501 GGTA AACAC AGATTATTTT CCCATTCTCA TTCGGCATT TTTCTGTACG
551 TATCATTTT TAGACGTATT TTTAGTCGAT TTGCCTTTTC CCGCATACCA
601 CGGCGCGGG TCGTCGGGCA GTCCGTCGAT AAAGGCAAGG TTATTGCCTT
651 CGCCCTGCAC ATCGGGAACA TTCCCCAAA AATCATAGCC GTCATCGGGC
701 AACTCGTCGG TTTGATACC CGTCCAACG CCGAATCCGC GTAA

```

This corresponds to the amino acid sequence <SEQ ID 258; ORF 082.a>:

```

a082.pep
  1 MWLLKLPVA K TALSPKRRR NTAANISFTV VLPPEVPVIPN TNGFTFSRHA
 51 FANICNAVSV S SFTFNAPSIA TQSSRETTTA AAPAANTSST KSCASNRP
101 NAKNTSPARM SRLSVTMRDT GLLSDGIGSL RAWQMKFRSS GFIFTFVNIR
151 AADTSVAADF FIACFAVVKH RLFSSHSAF FLYVSFFRRI FSRAFSRIP
201 RRGVVGQSV D KGKVI AFALH IGNIPPKIIA VIGQLVGFD T RPTAESAX

```

m082/a082 95.5% identity over a 247 aa overlap

	10	20	30	40	50	60
m082.pep	MXLLKLPVANTASSPKRRRNTAASISFTVVLPPPEVMPNTNGFTFSRHAFAVCNAASV					
a082	MWLLKLPVAKTALSPKRRRNTAANISFTVVLPPPEVPIPTNGFTFSRHAFAFANICNAVSV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m082.pep	SSTFNAPSIAAQSSRETTTAAAPAANTSSTKSCASNRS PANAKNTSPARMSRLSVTMRDT					
a082	SSTFNAPSIAATQSSRETTTAAAPAANTSSTKSCASNRP PANAKNTSPARMSRLSVTMRDT					
	70	80	90	100	110	120
	130	140	150	160	170	180
m082.pep	GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSXF					
a082	GLLSDGIGSLRAWQMKFRSSGFIFTFVNIRAADTSVAADFFIACFAVVKHRLFSHSHSAF					
	130	140	150	160	170	180
	190	200	210	220	230	240
m082.pep	FLYVSFFRRIFSRAFSRIPRRGVVGQSV DKGKVI AFALHIGNIPPKIIAVIGQLVGFD T					
a082	FLYVSFFRRIFSRAFSRIPRRGVVGQSV DKGKVI AFALHIGNIPPKIIAVIGQLVGFD T					
	190	200	210	220	230	240
m082.pep	RPTAESAX					
a082	RPTAESAX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 259>:

```

g084.seq
  1 ATGAAacaAT CCGccccgaat aAAAAATATG GATCAGACAT TAAAAAATAc
 51 attgggcatt tGCGCGcttt tagcctTTTG TTTTggcgcG gccatCGCAT
101 CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTTATTG TGCCGTGGGC
151 GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGCG GCTTCCCGCG
201 CGTTTCTTCA GTTGTTTTAC TGATTTACGT CGGCACAACC GCCCTATATT
251 TGCCGGTCGG CTGGCTGTAT GGTGCGCCTT CTTATCAGAT AGTCGGTTTCG
301 ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTCTG GCAATCTTCC
351 CGGGTCGCTT TATTTTGTGC AGGCATTATT TTTCAATTTT GGCTTGACAG

```

q084 . pep

```

1  MKQSARIKNN DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
51  ALASVVFLLL LARGFPRVSS VVLLTYVGGT ALYLPVGGWLY GAPS YQIVGS
101 ILESNPAEAR EFVGNLPGSL YFVQALFFDI GLTWVKYCVS VGVFADVKNY
151 KRRSKIWLT LLTLNLSCAV MEKTAGDKDW REPDAGLLLN IFDLYYDLAF
201 RAGTICROAR PHFGSSKKS V NMAYPPTCAQ V*

```

m084 . seg

1	ATGAAACAAT	CGCCCGAAT	AAAa .ATATG	AATCAGACAT	TACTTTATAC
51	ATTGGGCATT	TGCGCGCTTT	TAACCTTTnn	nnnnnnnnnn	nnnnnnnnnn
101	nnnnnnTATCA	CCnGAATAT	GAATACGGCT	ACCGTTATTC	TGCCGTGGGT
151	GCTTTGGCTT	CGGTTGTATT	TTTATTATTA	TTGGCAGCG	GTTTCCCGCT
201	CGTTTCTTCA	GTTGTTTAC	TGATTACGT	CGGCACAACC	CGCCTATATT
251	TGCCGGTCGG	CTGGCTGTAT	GGTGCGCCGT	CTTATCAGAT	AGTCGGTTCG
301	ATATTGGAAA	GCAATCCTGC	CGAGGCGCGT	GAATTGTGCG	GCAATCTTCC
351	CGGGTCGCTT	TATTTTGTGC	AGGCATTATT	TTTCATTTT	GGCTTGACAG
401	TTTGGAAATA	TGTGTATCG	GCGGGGTAT	TGTCTGACGT	AAAAACTAT
451	AAACGCGCGA	GCAAAATATG	GCTGACTATA	TTATTGACTT	TAAATTTGTC
501	CTGCGCGGTG	ATGGATAAAA	TCGCCAGCGA	TAAAGATTTG	CGAGAACCTG
551	ATGCCGCGCT	GTTGTTGAAT	ATTTTTCAGC	TGTATTACGA	TTTGGCT .TC
601	CGCGCCGCGA	CAATATGCCG	CCAAGCGCGC	CCACATTTTG	GAAGCAGCAA
651	AAAAAGCGCT	AACATGGCAT	ATCGCTCATG	TTCGCCCAA	GATAAA

m084.pcp

```

1  MKQSARIKXM  NQTLLEYTLGI  CALLTFXXXX  XXXXXYHPEY  EYGYRYSAVG
51  ALASVVFLLL  LARGFPRVSS  VVLLIYVGT  ALYLPVGWLY  GAPSQYIVGS
101 ILESNPAEAR  EFVGNLPGSL  YFQALFFDI  GLTWKVCVS  GGVFADVKNY
151 KRRSKIWLTI  LLTLILSCAV  MDKATSKDL  REPDAAGLLN  IFDLYYDLAX
201 RAGTICROAR  PHFGSSKSKS  NMAYSPSCAO  V*

```

Homology with a predicted ORF from *N. gonorrhoeae*

m084/q084

	10	20	30	40	50
m084 . pep	MKQSARIKXNMQTLTYLTGLICALLTF-----YHPEYEGYGRYSAVGALASVVFLLL				
	: :				
g084	MKQSARIKXNDQTLKNTLGLICALLAFCFGAAIASGYHLEYEGYGRYSAVGALASVVFLLL				
	10	20	30	40	50
	60	70	80	90	100
m084 . pep	LARGFPRVSSVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAAEAREFVGNLPGSL				
g084	LARGFPRVSSVLLIYVGTTALYLPVGWLYGAPSYQIVGSILESNPAAEAREFVGNLPGSL				
	70	80	90	100	110
	120	130	140	150	160
m084 . pep	YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWLTTILLTLILSCAVMDKIASDKDL				
g084	YFVQALFFIFGLTVWKYCVS VGVFADVKNYKRRSKIWLTTILLTLILSCAVMEKIAGDKDW				
	130	140	150	160	170
	180				

```
a084.seq
1 ATGAAACAAT CCGCCCGAAT AAAAAATATG GATCAGACAT TAAAAAATAC
51 ATTGGGCATT TGC CGCCTTTT TAACCTTTTG TTTTGGCGCG GCCATCGCAT
101 CAGGTTATCA CTTGGAATAT GAATACGGCT ACCGTTATTC TGCCGTGGGT
151 GCTTTGGCTT CGGTTGTATT TTTATTATTA TTGGCACGGG GTTTCGCCGG
201 CGTTTCTTCA GTTGTTTAC TGTATTACGT CGGCACAACC GCCCATATTT
251 TGCCGGTCGG CTGGCTGATT GGTGCGCCGT CTTATCAGAT AGTCGTTTCG
301 ATATTGGAAA GCAATCCTGC CGAGGCGCGT GAATTTGTGC GCAATCTTCC
351 CGGGTGCCTT TATTTTGTGC AGGCATTATT TTTCAATTTT GGCTTGACAG
401 TTTGGAGATA TTGTGTATCG GGGGGGGTAT TTGCTGACGT AAAAAAGTAT
451 AAACGCGCGCA GCAAAATATG GCTGACTATA TTATTGACTT GTATTTGTGC
501 CTGCGCGGGT ATGGATAAAA TCGCCAGCGA TAAAGATTTG CGAGAACCTG
551 ATGCGCGGCT GTTGTTGAAT ATTTTTCAGC TGTATTACGA TTTGGCTTCC
601 .GCGCCGGCA CAATATGCCG CCAAGCGCGC CCACATTTTG GAAGCAGCAA
651 AAAAAGCGTC AACATGGCAT ATCCGTATG TTGCGCCCAA GTATAA
```

```
a084.pep
1  MKQSARIKNN DQTLKNTLGI CALLAFCFGA AIASGYHLEY EYGYRYSAVG
51  ALASVVELLL LARGFRVSS VLLIYVGTG ALYLPVGWLY GAFSYQIVGS
101 ILESNPPEAR EFVGNLPGSL YFVQALFFIF GLTWVRYCVS GGVFADVKNY
151 KRRSKIWLTI LLTLILSCAV MKDIASDKDL REPDAGLLLN IFDLYYDLAS
201 XAGTICROAR PHFGSSKKSX NMAYPSCCAQ V*
```

	10	20	30	40	50	60
m084.pep	MKQSARIKXNMQTLTYLTGICALLTFXXXXXXXYHPEYEGYGRYS	AVGALASVVFLL				
a084	MKQSARIKXNMQTLKNTLGICALLAFCFGAATASGYHLEYEGYGRYS	AVGALASVVFLL				
	10	20	30	40	50	60
m084.pep	LARGFPRVSSVLLIYVGT	TALYLPVGWLYGAPSYQIVGS	ILESNPAEAREFVGNLPGSL			
a084	LARGFPRVSSVLLIYVGT	TALYLPVGWLYGAPSYQIVGS	ILESNPAEAREFVGNLPGSL			
	70	80	90	100	110	120
m084.pep	YFVQALFFIFGLTVWKYCVSGGVFADVKNYKRRSKIWL	TILLTLILSCAVMDKIASDKDL				
a084	YFVQALFFIFGLTVWRYCVSGGVFADVKNYKRRSKIWL	TILLTLILSCAVMDKIASDKDL				
	130	140	150	160	170	180
m084.pep	REP	DAGLLNIFDLYDLAXRAGTICRQARPHFGSSKKS	SVNMAYPSCCAQVX			
a084	REP	DAGLLNIFDLYDLASXAGTICRQARPHFGSSKKS	SVNMAYPSCCAQVX			
	190	200	210	220	230	

g085.seq
1 ATGGGGCAAAG GGCAGGACTT CACGCCCCCTG CGCGACGCGT TGAAAGATAA

```

51  GGCAAAAGGC GTGTTCTGA TCGCGCTCGA TCGCGCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AACCTGACCG ACTGCGTCAC TTTGGAAGAG
151 GCGGTTTCTAGA CCGCATACGC CCAAGCCGAA GCGGGCGATA TTGTCTTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
251 CGGAAGTGTt tatCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 266; ORF 085.ng>:

g085.pep

```

1  MGKGQDFTPL RDALKDKAKG VFLIGVDAPO IRRDLGCGL NLTDCVTLEE
51 AVQTAYAQAe AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 267>:

m085.seq

```

1  ATGGGTAAAG GGCAGGACTT CACGCCCTG CGCGATGCAC TGGTAGGCAA
51  GGCAAAAGGC GTGTTCTGA TTGGTGTCTGA TCGCGCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGGCTTG AATATGACCG ACTGCGCCAC TTTGGGAGAA
151 GCCGTTTCTAGA CCGCATATGC CCAAGCCGAA GCAGGCGATA TTGTGTTGCT
201 CAGCCCCGCC TCGCGAGCT TTGATATGTT CAAAGGCTAC GCGCACCGTT
251 CGGAAGTGTt TATCGAAGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 268; ORF 085>:

m085.pep

```

1  MGKGQDFTPL RDALVGKAKG VFLIGVDAPO IRRDLGCGL NMTDCATLGE
51 AVQTAYAQAe AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 085 shows 94.7% identity over a 94 aa overlap with a predicted ORF (ORF 085.ng) from *N. gonorrhoeae*:

m085/g085

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGLNMTDCATLGEAVQTAYAQAe					
g085	MGKGQDFTPLRDALKDKAKGVFLIGVDAPOIRRDLGCGLNLTDCVTLEEAVQTAYAQAe					
	10	20	30	40	50	60
	70	80	90			
m085.pep	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
g085	AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX					
	70	80	90			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 269>:

a085.seq

```

1  ATGGGCAAAG GGCAGGACTT CACGCCCTG CGCGACGCGC TTGCCGGCAA
51  GGCAAAAGGC GTGTTCTGA TCGGTGTCTGA TCGCGCGCAA ATCCGCCGCG
101 ATTTGGACGG CTGCGATCTG AATATGACCG ACTGCGCCAC TTTGGAAGAA
151 GCGGTTTCTAGA AGGCATATGC CCAAGCCGAA GCGGGCGATA TCGTGCTGCT
201 CAGCCCCGCC TCGCGAGTT TCGATATGTT TAAAGGCTAC GCGCACCGTT
251 CGGAAGTGTt TATCGGGGCG TTTAAGGCTT TGTGA

```

This corresponds to the amino acid sequence <SEQ ID 270; ORF 085.a>:

a085.pep

```

1  MGKGQDFTPL RDALAGKAKG VFLIGVDAPO IRRDLGCDL NMTDCATLEE
51 AVQKAYAQAe AGDIVLLSPA CASFDMFKGY AHRSEVFIEA FKAL*

```

m085/a085 94.7% identity over a 94 aa overlap

	10	20	30	40	50	60
m085.pep	MGKGQDFTPLRDALVGKAKGVFLIGVDAPOIRRDLGCGLNMTDCATLGEAVQTAYAQAe					

274

```

a085      |||||:||||| ||||| ||||| ||||| |||||
          10      20      30      40      50      60
          70      80      90
m085.pep  AGDIVLLSPACASFDMFKGYAHRSEVFIEAFKALX
          ||||| ||||| ||||| ||||| |||||
a085      AGDIVLLSPACASFDMFKGYAHRSEVFIGAFKALX
          70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 271>:

```

g086.seq
1  ATGGTGGTGC TGATGACGGC GTTCGGCCTG CTGATGATT ATTCGGCTTC
51  TGTGTATTTC GCATCGAAGG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
101 GGCAGGCGGG GTTCGTCGTT GCCGGCCTTA TAGCGAGCGG TTTTATATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTGTGCCGT GGATTTTTCG
201 CTTATCCGGC CTGTTGCTGG TAGCCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTGGGTC CGTTGAATT CCAGCCGACC
301 GAGCTGTTCA AGCTGGCAGT CATCCTTTAT TTGGCAAGCC TGTTACGCG
351 CCGTGAAGAA GTGTTGCGCA GCATGGAAG TTTGGGTTGG CAGTCGATT
401 GGCAGGGGAC GGCCAACCTG ATTATGTCCG CCACCAATCC GCAGGCACGT
451 CGTGAAACAT TAGAAATGTA CGGCGGTTTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG
551 GTTCGTTTGT CGTCATTACC GTCAATACCG TTGGAATGCT GTTCTGGCA
601 GGATTGCCGT GGAAATATTT TTTTGTCTGT GTAGGCAGCG TCTTGGGTGG
651 GATGGTGCTG ATGATTACCG CCGCTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTGGGA CCGGTGGAAG GACCCGACAG GTGCCGCTA CCAGCTTACC
751 CACTCTCTGA TGGCAATCGG GCGCGGAGAG TGGTTCGGTA TGGGTTTGGG
801 TGGGAGTTTG AGCAAACGCG GCTTCTGCC GGAAGCGCAT ACCGATTTA
851 TTTTGGCAT CATCGCTGAA GAATTCGGCT TCTTCGGGAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGGTGGG GCGTTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGtttG CTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTGGAT CCGTATCCAA AGTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA CCAAAGGTCT GACGctgCcg tGATGTCTT ATGGcggTTC
1101 GTCAGTCTTT TTCATGCTGA TCAGCATGAT GCTGCTGTT CGTATCGATT
1151 ATGAAACCG CCAGAAATG CCGGTTACC GGGTGGAGTA AA

```

This corresponds to the amino acid sequence <SEQ ID 272; ORF 086.ng>:

```

g086.pep
1  MVLMTAFGL LMIYSASVYL ASKEGGDQFF YLTRQAGFVV AGLIASGFLW
51  FLCRMRTWRR LVPWIFALSG LLLVAVLIAG REINGATRWI PLGPLNFQPT
101 ELFKLAVILY LASLFTRREE VLRSMESLGW QSIWRGTANL IMSATNPQAR
151 RETLEMYGRF RAILPIMLV AFGLVLIMVQ PDFGSFVVIT VITVGMFLFLA
201 GLPWKYFFVL VGSVLGGMVL MITAAPYRVQ RVVAFDPWK DPQGAGYQLT
251 HSLMAIGRGE WFGMGLGASL SKRGFLPEAH TDFIFAIIE EFGFFGMCVL
301 IFCYGLVVR AFSIGKQSRD LGLTFNAYIA SGIGIWIQI SFFNIGVNIG
351 ALPTKGLTLP LMSYGGSSVF FMLISMLLL RIDYENRQKM RGYRVE*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 273>:

```

m086.seq
1  ATGGTGGTGC TGATGACGGC GTTCAGCCTG CTGATGATT ATTCGGCTTC
51  TGTGTATTTC GCATCAAAAG AAGGCGGCGA TCAGTTTTTC TATTTGACCA
101 GACAGGCGGG GTTCGTCGTT GCCGGCTTGA TAGCGAGCGG TTTGTTATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTGTGCCGT GGATTTTTCG
201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCGGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTGGGTC CGTTGAATT CCAGCCGACC
301 GAGCTGTTCA AGCTGGCGGT CATCCTTTAT TTGGCAAGCC TGTTACGCG
351 CCGTGAAGAA GTGTTGcGCA GCATGGAAG TTTGGGTTGG CAGTCGATT
401 GGCAGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGrCACGT
451 CGTGAAaACAT TAGAAATGTA CGGCCGTwTC CGGGCGATCA TCCTGCCGAT
501 TATGCTGGTG GCGTTCGGTT TGGTGCTGAT AATGGTACAG CCGGATTTCG

```


m086.pcp

m086/g086

	10	20	30	40	50	60
m086.pep	MVVLMTAFSLLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
g086	MVVLMTAFGLLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLFWFLCRMRTWRR					
	10	20	30	40	50	60
m086.pep	LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTL FKLAVILYLASLFTRREE					
g086	LVPWIFALSGLLLVAVLIAGREINGATRWIPLGPLNFQPTL FKLAVILYLASLFTRREE					
	70	80	90	100	110	120
m086.pep	VLRSMESLQWQSIWRGTANLIMSATNPQXRRETLEMYGRXRAILPIMLVAFGLVLIMVQ					
g086	VLRSMESLQWQSIWRGTANLIMSATNPQARRETLEMYGRFRAILPIMLVAFGLVLIMVQ					
	130	140	150	160	170	180
m086.pep	PDFGSFVVITTVIAGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
g086	PDFGSFVVITTVIAGMLFLAGLPWKYFFVLVGSVLGGMVLMITAAPYRVQRVVAFLDPWK					
	190	200	210	220	230	240
m086.pep	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFFGFGMCVL					
g086	DPQGAGYQLTHSLMAIGRGEWFGMGLGASLSKRGFLPEAHTDFIFAIIAEEFFGFGMCVL					
	250	260	270	280	290	300
m086.pep						
g086						
	310	320	330	340	350	360

m086.pep
|FCYGLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGXQSFFNIGVNIGALPXXKGLTXP
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| ||
g086
IFCYGLVVRAFSIGKQSRDLGLTFNAYIASGIGIWIGIQSFFNIGVNIGALPTKGLTLP
 310 320 330 340 350 360

m086.pep
 370 380 390
XMSXGGSSVF FMLISM LLLXRIDYENRRKMGRGYRVEX
|| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| ||
g086
LMSYGGSSVF FMLISM LLLLRIDYENRQKMRGYRVEX
 370 380 390

```
a086.seq
1   ATGGTGGTGC TGATGACGGC GTTACAGCCTG CTGATGATTT ATTCGGCTTC
51  TGTGTATTTG GCATCAAAAG AAGCGGCGCA TCAGTTTTTT TATTTGACCA
101 GACAGGCGGG GCTCGTCGTT GCGCGCTTGA TAGCGAGCGG TTTGTTATGG
151 TTTCTTTGCA GGATGAGGAC ATGGCGGCGG CTTGTGCCGT GGATTTTTGC
201 CCTATCCGGC CTGTTGCTGG TAGTCGTATT GATTGCCCGG CGCGAAATCA
251 ATGGCGCGAC CCGTTGGATA CCTTTGGGTC CGTTGAATTT CCAGCCGACC
301 GAGCTGTAAC AGCTGGCGGT CATCTTTAT TTGGCAAGCC TGTTCCAGCG
351 CCGTGAAAGAA GTGTTGCGCA GCATGGAAAG TTTGGGTTGG TAGTCGATTT
401 GCGGGGGGAC GGCCAATCTG ATCATGTCCG CCACCAATCC GCAGGCACGT
451 CGTGAACATC TAGAAATGTA CGGCCGTTTC CGGGCGATCA TCCTCGCCAT
501 TATGCTGGTG GTGCTCGGTT TGGTGCTGAT AATGGTACAG CGGGATTTGC
551 GTTCGTTGTG CGTCATATACC GTCATTGCCG TTGGAATGCT GTTTTTGGCA
601 GGATTGCCGT GGAAATATTT TTTGTCCTG GTAGGCAGCG TCTTGGGCGG
651 GATGGTGCTG ATGATATACC CGCGTCCCTA CCGTGTGCAG CGGGTAGTGG
701 CATTTTTGGA CCCGTGGAAA GACCCGAGG GTGCCGGCTA CCAGCTTACC
751 CACTCTCTGA TGCCAATCGG GCGCGGAGAG TGGTTTCGTA TGGGTTTGGG
801 TGCGAGTTTG AGCAAAACGC GCTTTCGTCC GGAAGCGCAT ACCGATTTTA
851 TTTTGTCCAT CATCGCCGAA GAATTCGGTT TCCTTCGGTAT GTGCGTGCTG
901 ATATTCTGTT ACGGCTGGCT GGTGTCGCGG GCGTTTTCCA TCGGCAAGCA
951 GTCGCGCGAT TTGGGTTTGA CTTTCAACGC CTATATCGCT TCGGGTATCG
1001 GCATTTGGAT CGGTATCCAA AGTTTCTTCA ATATCGGTGT GAACATCGGT
1051 GCTTTGCCGA CCAAAGGTCT GACGCTGCCG TTGATGTCTT ATGCGCGTTC
1101 GTCAGCTCTT TTCATGCTGA TCAGCATGAT CGTGCTGTGG CGTATAGATT
1151 ATGAAACACG CCGGAAAATG CGCGGTTACC GGGTGGAGTA A
```

a086.pap

1	MVVLMTAFSL	LMIYSASVYL	ASKEGGDQFF	YLTRQAGFVV	AGLIASGLLW
51	FLCRMRTWRR	LVPWIFALSG	LLLUVVLIAG	REINGATRWI	PLGLNFQPT
101	ELFKLAVILY	LASLFTFREE	VLRSMESLWG	QSWRGRTANL	IMSATNPQAR
151	RETLEMYGRF	RAIILPIMLV	AFGLVLIMVQ	PDFGSEFVIT	VIAVGMLFLA
201	GLPWKYFFVL	VGSVLGGMVL	MITAAPYRVQ	RVVAFLLDPWK	DPQAGYQLT
251	HSLMAIGRGE	WFGMGLGASL	SKRGLFPEAH	TDFIFAIIAE	EPGFFGMCVL
301	IFCYGWLVRV	AFSIGKQSRD	KLGFENAYIA	SGIGIWIGIQ	SFFNIGCVNIG
351	ALPTKGLTLP	LMSYGGSSVF	FMLISMMLLL	RIDYENRRRK	RGYRVE*

	10	20	30	40	50	60
m086.pep	MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
a086	MVVLMTAFSLLMIYSASVYLASKEGGDQFFYLTRQAGFVVAGLIASGLLWFLCRMRTWRR					
	10	20	30	40	50	60
	70	80	90	100	110	120
m086.pep	LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTLEFKLAVILYLASLFTTRREE					
a086	LVPWIFALSGLLLVVVLIAGREINGATRWIPLGPLNFQPTLEFKLAVILYLASLFTTRREE					

277

	70	80	90	100	110	120
m086.pep	130	140	150	160	170	180
a086	130	140	150	160	170	180
m086.pep	190	200	210	220	230	240
a086	190	200	210	220	230	240
m086.pep	250	260	270	280	290	300
a086	250	260	270	280	290	300
m086.pep	310	320	330	340	350	360
a086	310	320	330	340	350	360
m086.pep	370	380	390			
a086	370	380	390			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 277>:

```

g087.seq
1  ATGGGCGGTA AAACCTTTAT GCTGATGGCG GCGGGAACGG GCGGACACAT
51  TTTCCAGCT CTGGCTGTGG CGGATTCATT GCGCGTGCGC GGTCAATCATG
101 TAATTGGCT GGGCAGCAAG GATTGATGG AAGAGCGCAT CGTGCCGCAA
151 TACGGCATA CTTTGAAAC GCTGGCGATT AAAGGAATAC GCGCAACCGG
201 CATCAAACGC AAGCTGATGC TTCCGTTTAC TCTGTACAAA ACCGTCCGCG
251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCGT CATCGGCTTC
301 GCGGTTTGT TTACCTTTCC CGGCGGTCTG GCGGCGAAAC TCTTGGGCGT
351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGCTTG TCCAACCGCC
401 AccTGTGCGC ctGGGCGAAA CGGGTGTGTG ACGCTTTTCC GAAAGCGTTC
451 AGCCACGAAG GCGGTTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGCGAAGGC CGTCTGAAAA
551 TTTTGGTGGT CGGCGGCAGT TTGGGTGCGG ACGTTTTGAA CAAAACCGTA
601 CCGCAGGCGT TGGCACTGCT GCCTGAAGAG GTGCGCCCGC AGATGTACCA
651 CCAGTCGGGG CGTAACAAGC TGGGCAATCT TCAGGCGGAT TATGACGCGT
701 TGGGCGTGAA AGCGGAATGC GTGGAATTTA TTACCGACAT GGTGTCCGCC
751 TACCGTGATG CCGATTGGT GATTTGCCGT GCCGGCGCGC TGACGATTGC
801 CGAGTTGACG GCGGCGGGGC TGGGCGCGTT GTTAGTGCCG TATCCTCACG
851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTCATGGT GCAGGCAGAA
901 GCGGGGCTGC TGTGCGGCA AACCAGTTG ACGGCGGAAA AACTCGCCGA
951 AATCCTCGGC AGCCTCAACC GCGAAAAATG CCTCAAATGG GCGGAAAAACG
1001 CCCGTACGTT GGCATTGCCG CACAGCGCGG ATGACGTTGC CGAAGCCGCG
1051 ATTGCGTGTG CGGCGTAAA

```

This corresponds to the amino acid sequence <SEQ ID 278; ORF 087.ng>:

```

g087.pep
1  MGGKTFMLMA GGTGGHIFPA LAVADSLRVR GHVIVLWLSK DSMEERIVPQ
51  YGIRLETIAI KGIRNGIKR KLMLPFTLYK TVREAQRIIR KHRVECVIGF
101 GGFVTFPGGL AAKLLGVPIV IHEQNAVAGL SNRHLRSWAK RVLVAFPKAF
151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGG LGADVLNKTV

```

201 PQALALLPEE VRPQMYHQSG RNKLGNLQAD YDALGVKAEC VEFITDMVSA
 251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
 301 AGLLLPQTQL TAEKLAIEILG SLNREKCLKW AENARTLALP HSADDVAEAA
 351 IACAA*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 279>:

m087.seq
 1 ATGGGCGGTA AAACCTTTAT GCTGAWkkCG GCGGAACGG GCGGACATAT
 51 TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
 101 TGATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGTAT CGTGCCGCAA
 151 TACGGCATAAC GCTTGGAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
 201 CATCAAACGC AAACCTGATGC TGCCGGTTAC TTTGTATCAA ACCGTCCGCG
 251 AAGCGCAGCG GATTATCCGC AAACACCGTG TCGAGTGCCT CATCGGCTTC
 301 GCGCGCTTCG TTACCTTCCC CGGCGGTTTG GCGGCGAAGC TATTArGCGT
 351 GCCGATTGTG ATTCACGAGC AAAACGCCGT GGCAGGTTTG TCCAACCGCC
 401 ACCTGTGCGC CTGGGCGAAG CGGGTGTTGT ACGCTTTTCC GAAAGCGTTC
 451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
 501 CCTGCCCGTG CCTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
 551 TTTTGGTGGT CGGCGGCACT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA
 601 CCGCATGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC ATATGTACCA
 651 CCAATCGGGA CGGGGCAAGC TGGGCATCTT GCAGGCGGnnn nnnnnnnnnn
 701 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 751 nnnGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG CGGTTGACGA
 801 TCACCAAACC GCCAACGCGC GTTTATGGT GCAGGCGGAG GCGGGATTGC
 851 TGTGTCGCGA AACCAGTTG ACGGCGGAAA AACTCGCCGA GATTCTCGGC
 901 GGCTTAAACC GCGAAAAATG CCTCAAATGG GCAGAAAACG CCCGTACGTT
 951 GGCAGTCCG CACAGTCCG ACGACGTGGC GGAAGCCGCG ATTGCGTGTG
 1001 CGGCGTAA

This corresponds to the amino acid sequence <SEQ ID 280; ORF 087>:

m087.pep
 1 MGGKTFMLXX GGTGGHIFPA LAVADSLRAR GHVHVLGSK DSMEERIVPQ
 51 YGIRLETLAI KGVRGNGIKR KLMLPVTLYQ TVREAQRIIR KHRVECVIGF
 101 GGFVTFPGGL AAKLLXVPIV IHEQNAVAGL SNRHLRWAK RVLVAFPKAF
 151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGS LGADVNLKTV
 201 PHALALLPDN ARPHMYHQSG RGKLILQAX XXXXXXXXXX XXXXXXXXXX
 251 XAGLGALLVP YPHAVDDHQT ANARFMVQAE AGLLLPQTQL TAEKLAIEILG
 301 GLNREKCLKW AENARTLALP HSADDVAEAA IACAA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 087 shows 83.9% identity over a 355 aa overlap with a predicted ORF (ORF 087.ng) from *N. gonorrhoeae*:

m087/g087

	10	20	30	40	50	60
m087.pep	MGGKTFMLXXGGTGGHIFPALAVADSLRARGHVVHVLGSKDSMEERIVPQY	GIRLETLAI				
g087	MGGKTFMLMAGGTGGHIFPALAVADSLRVRGHVVHVLGSKDSMEERIVPQY	GIRLETLAI				
	10	20	30	40	50	60
m087.pep	KGVRGNGIKRKLMLPVTLYQTVREAQRIIRK	KHRVECVIGFGGFVTFPGGLAAKLLXVPIV				
g087	KGIRGNGIKRKLMLPFTLYKTVREAQRIIRK	KHRVECVIGFGGFVTFPGGLAAKLLGVPIV				
	70	80	90	100	110	120
m087.pep	IHEQNAVAGLSNRHLRWAKRVLVAFPKAFS	HEGGLVGNPVRADISNLPVPAERFQREG				
g087	IHEQNAVAGLSNRHLRWAKRVLVAFPKAFS	HEGGLVGNPVRADISNLPVPAERFQREG				
	130	140	150	160	170	180

279

```

              190      200      210      220      229
m087.pep    RLKILVVGGSLGADVLNKTVPHALALLPDNARPHMYHQSGRGKGLQLQA-----
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
g087        RLKILVVGGSLGADVLNKTVPQALALLPEEVRPQMYHQSGRNKLGNLQADYDALGVKAEC
              190      200      210      220      230      240

              230      240      250
m087.pep    -----AGLGALLVPYPHAVDDHQTANARFMVQAE
              |||||:|||||:|||||:|||||:|||||:|||||
g087        VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE
              250      260      270      280      290      300

              260      270      280      290      300      310
m087.pep    AGLLLPQTQLTAEKLAELGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
g087        AGLLLPQTQLTAEKLAELGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX
              310      320      330      340      350

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 281>:

```

a087.seq
1   ATGGGCGGTA AAACCTTTAT GCTGATGGCG GCGGGAACGG GCGGACATAT
51  TTTCCCCGCG CTGGCGGTGG CGGATTCATT GCGCGCGCGC GGCCATCATG
101 TAATTTGGCT GGGCAGCAAG GATTCGATGG AAGAGCGCAT CGTGCCGCAA
151 TACGACATCC TGCTCGAAAC GCTGGCGATT AAAGGCGTGC GCGGCAACGG
201 CATCAAACGC AAGCTGATGC TGCCGTTTAC TTTGTATCAA ACTGTCCGCG
251 AAGCGCAGCA GATTATCCGC AAACACCGTG TCGAGTGCCT CATCGGCTTC
301 GCGCGCTTCG TTACCTTTCC CGGCGGTTTG GCGGCGAAGT TATTAGCGCT
351 GCCGATTGTG ATTACAGAGC AAAACGCCGT GGCAGGTTTG TCCAACGCC
401 ACCTGTCGCG CTGGGCGAAG CGGCTGTGTG ACGCTTTTCC GAAAGCGTTC
451 AGCCACGAAG GCGGCTTGGT CGGCAACCCC GTCCGCGCCG ATATTAGCAA
501 CCTGCCCGTG CTGCCGAAC GCTTCCAAGG GCGTGAAGGC CGTCTGAAAA
551 TTTTGGTGGT CGGCGGCAGT TTGGGCGCGG ACGTTTTGAA CAAAACCGTA
601 CCGCAGGCAT TGGCTTTGCT GCCCGACAAT GCGCGTCCGC AGATGTACCA
651 CCAATCGGGA CGGGCAAGC TGGGCAGCTT GCAGGCGGAT TACGACGCGC
701 TGGGCGTGCA AGCGGAATGC GTGGAATTTA TTACCGATAT GGTGTCCGCC
751 TACCGCGATG CCGATTGGT GATTGCGCGT GCCGCGCGC TGACGATTGC
801 CGAGTTGACG GCGGCGGGAT TGGGTGCGTT GTTAGTGCCG TATCCTCACG
851 CCGTTGATGA CCATCAAACC GCCAACGCGC GTTTTATGGT GCAGGCGGAG
901 GCGGGATTGC TGTGCGGCA AACCCAGTTG ACGGCGGAAA AACTCGCCGA
951 GATTCTCGCG GGCTTAAACC GCGAAAATG CCTCAAATG GCAGAAAACG
1001 CCCGTACGTT GGCAGTCCG CACAGTGCGG ACGACGTTGC CGAAGCCGCG
1051 ATTGCGTGTG CGGCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 282; ORF 087.a>:

```

a087.pep
1   MGGKTFMLMA GGTGGHIFPA LAVADSLRAR GHVVIWLGSK DSMEERIVPQ
51  YDILLETLAI KVRGNGIKR KLMLPFTLYQ TVREAQQIIR KHRVECVIGF
101 GGFVTFPGLL AAKLLGVPIV IHEQNAVAGL SNRHLRWAK RVLVAFPKAF
151 SHEGGLVGNP VRADISNLPV PAERFQREG RLKILVVGGG LGADVLNKTVP
201 PQALALLPDN ARPQMYHQSG RGKLGSLQAD YDALGVQAE VEFITDMVSA
251 YRDADLVICR AGALTIAELT AAGLGALLVP YPHAVDDHQT ANARFMVQAE
301 AGLLLPQTQL TAEKLAELG GLNREKCLKW AENARTLALP HSADDVAEAA
351 IACAA*

```

m087/a087 85.4% identity over a 355 aa overlap

```

              10      20      30      40      50      60
m087.pep    MGGKTFMLXXGGTGGHIFPALAVADSLRARGHHVVIWLGSKDSMEERIVPQYIGIRLETLAI
              |||||:|||||:|||||:|||||:|||||:|||||:|||||
a087        MGGKTFMLMAGGTGGHIFPALAVADSLRARGHHVVIWLGSKDSMEERIVPQYDILLETLAI
              10      20      30      40      50      60

```

280

	70	80	90	100	110	120
m087.pep	KGVRRNGIKRKLMLPVTLYQTVREAQRIIRKHRVECVIGFGGFVTFPGGLAAKLLXVP					
a087	KGVRRNGIKRKLMLPFTLYQTVREAQQIIRKHRVECVIGFGGFVTFPGGLAAKLLGVPIV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m087.pep	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
a087	IHEQNAVAGLSNRHLSRWAKRVLYAFPKAFSHEGGLVGNPVRADISNLPVPAERFQGREG					
	130	140	150	160	170	180
	190	200	210	220	230	240
m087.pep	RLKILVVGSLGADVLNKTVPHALALLPDNARPHMYHQSGRGKLGILQAXXXXXXXXXXX					
a087	RLKILVVGSLGADVLNKTVPQALALLPDNARPMYHQSGRGKLGSLQADYDALGVQAE					
	190	200	210	220	230	240
	250	260	270	280	290	300
m087.pep	XX-----XXXXXXXXXAGLGALLVPYPHAVDDHQTANARFMVQAE					
a087	VEFITDMVSAYRDADLVICRAGALTIAELTAAGLGALLVPYPHAVDDHQTANARFMVQAE					
	250	260	270	280	290	300
	290	300	310	320	330	
m087.pep	AGLLLPQTQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
a087	AGLLLPQTQLTAEKLAELGGLNREKCLKWAENARTLALPHSADDVAEAAIACAAX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 283>:

g088.seq

```

1  ATGTTTATAT GGCTCGCACA TTTCAGCAAC TGGTTAACCG GTCTGAATAT
51  TTTTCAATAC ACCACATTCC GCGCCGTTAT GCGCGCGTTG ACCGCCTTGG
101 CGTTTTCCTT GATGTTCCGC CCGTGGACGA TACGCAGGCT GACCGCGCTC
151 AAATGCGGGC AGGCAGTGCG TACCGACGCG CCGCAAACCC ACCTCGTCAA
201 AAACGGCAGC CCGACGATGG GCGGTTCCGT GATTCTGACC GCCATTACCG
251 TGTCCACCCT GTTGTGGGGC AACTGGGCGA ACCCGTATAT CTGGATTCTC
301 TTGGGCGTAC TGCTTGCCAC CCGTGCGCTC GGTTTTTACG ACGACTGGCG
351 CAAAGTCGTT TATAAAGACC CCAACGGCGT GTCCGCCAAA TTCAAATGG
401 TGTGGCAGTC AAGCGTTGCC GTTatcgCG GTttggcaTT GTTTTACctt
451 gCcgCAATT CCGCCAACAA TATTTTGATT GTCCCGtttT TCAAACAAAT
501 CGCCCTGCCG CTGGGCGTGG TCGGCTTttt gGtgttgTCT TACCTGACCA
551 TCGTCGGCAC ATCCAACGCC GTCAACCTCA CcgaCGGCTT GGACGGCCTT
601 GCCGCcttcc cgttegtcct cgttgccgcC GGGCTCGCCA ttttcgccTA
651 CGTCAGCGGA CACTACCAAT TTTCCCAATA CCTCCAGCTT CCCTATGTCG
701 CCGGCGCGAA CGAAGTCGCT ATATTCTGCA CCGCCATGTG CCGCGCGTGC
751 CTCGGATTTT TGTGGTTCAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
801 TGTGCGCGCG CTGGCATTGG GTGCCGCGCT CGGTaccGtt gCCGTcaTcg
851 tCCGCCAAGA ATTTGTcctc gtcattatGG GCGGTCTGTT cgtcgtagaa
901 gccgtgTCCG TTATGCTTca tgtcggCTGG TACAAGAAAA Ccaaaaaacg
951 CATCTTcCTg acgGcaccga ttcacacacca ttaCCaactt cgatgCTGGa
1001 aagaaacgca agtcgctcgtc CGTtCTGGA TTAtTaccat cgtcgtggtt
1051 tTgatagggt tGagtaccct caAAattcgc ggaaactatg ccgTCCGAAC
1101 ACCTTTCAGA CCGCATTTGA ACGCGCAATA A

```

This corresponds to the amino acid sequence <SEQ ID 284; ORF 088.ng>:

g088.pep

```

1  MFLWLAHFSN WLTGLNIFQY TTFRAVMAAL TALAFSLMFG PWTIRRLTAL
51  KCGQAVRTDG PQTHLVKNGT PTMGGSLILT AITVSTLLWG NWNAPYIWL
101 LGVLLATGAL GFYDDWRKVY YKDPNGVSAK FKMVWQSSVA VIAGLALFYL
151 AANSANNILI VPFFKQIALP LGVVGFVLVS YLTIVGTSNA VNLTGDLGDL

```

201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC
 251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
 301 AVSVMLHVWG YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
 351 LIGLSTLKIR GNYAVRTPFR RHLNAQ*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 285>:

m088.seq
 1 ATGTTTTTAT GGCTCGCACA TTTCAGCAnC TGGTTAACCG GTCTGAATnn
 51 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 101 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 151 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 201 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 251 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 301 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 351 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 401 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 451 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn
 501 nnnnnnnnnn nnnGGCGTGG TCGGCTTTT GGTGTGTCT TACCTGACCA
 551 TCGTCGGCAC ATCCAATGCC GTCAACCTCA CCGACGGCTT GGACGGCCTT
 601 GCGACCTTCC CCGTCGTCCT CGTTGCCGCC GGCCTCGCCA TCTTCGCCTA
 651 TGCCAGCGGC CACTCACAAT TTGCCCAATA CCTGCAATTA CCTTACGTTG
 701 CCGGCGCAAA CGAAGTGGTG ATTTTCTGTA CCGCCATGTG CGGCGCGTGC
 751 CTCGGTTTCT TGTGGTTTAA CGCCTATCCC GCGCAAGTCT TTATGGGCGA
 801 TGTCGGTGCA TTGGCATTGG GTGCCCGGCT CGGTACCGTC GCCGTTATCG
 851 TCCGCCAAGA GTTTGTCCTC GTCATTATGG GCGGATTATT TGTCGTAGAA
 901 GCCGTATCCG TTATGCTTCA GGTGGCTGG TATAAGAAAA CCAAAAAACG
 951 CATCTTCTCG ATGGCGCCCA TCCATCACCA CTACGAACAA AAAGGCTGGA
 1001 AAGAAACCCA AGTCGTCGTC CGCTTTTGA TTATTACCAT CGTCTTGGTG
 1051 TTGATCGGTT TGAGTACCTT CAAAATCCGC TGAACCTATG CCGTCTGAAC
 1101 ATCTTTCAGA CGGCATTGA ACGCGCAATA A

1 MFLWLAHFSN WLTGLNIFQY TTFRVMAAL TALAFSLMFG PWTIRRLTAL
 51 KCGQAVRTDG PQTHLVKNGT PTMGGSILIT AITVSTLLWG NWANPYIWIL
 101 LGVLLATGAL GFYDDWRKVY YKDPNGVSAC FKMVWQSSVA VIAGLALFYL
 151 AANSANNILI VPFFKQIALP LGVVGFLVLS YLTIVGTSNA VNLTGDLGDL
 201 AAFPFVLVAA GLAIFAYVSG HYQFSQYLQL PYVAGANEVA IFCTAMCGAC
 251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
 301 AVSVMLHVWG YKKTKKRIFL TAPIHHHYQL RCWKETQVVV RFWIITIVVV
 351 LIGLSTLKIR GNYAVRTPFR RHLNAQ*

This corresponds to the amino acid sequence <SEQ ID 286; ORF 088>:

m088.pep
 1 MFLWLAHFSX WLTGLNXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
 51 XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
 101 XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX XXXXXXXXXXX
 151 XXXXXXXXXXX XXXXXXXXXXX XGVVGFLVLS YLTIVGTSNA VNLTGDLGDL
 201 ATFPFVLVAA GLAIFAYASG HSQFAQYLQL PYVAGANEVV IFCTAMCGAC
 251 LGFLWFNAYP AQVFMGDVGA LALGAALGTV AVIVRQEFVL VIMGGLFVVE
 301 AVSVMLQVGW YKKTKKRIFL MAPIHHHYEQ KGWKETQVVV RFWIITIVLV
 351 LIGLSTLKIR XTYAVXTSFR RHLNAQ*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 088 shows 91.7% identity over a 205 aa overlap with a predicted ORF (ORF 088.ng) from *N. gonorrhoeae*:

m088/g088

m088.pep				10	20	30																																																						
				G	V	V	G	F	L	V	L	S	Y	L	T	I	V	G	T	S	N	A	V	N	L	T	D	G	L	D	G	L	A																											
g088	I	A	G	L	A	L	F	Y	L	A	A	N	S	A	N	N	I	L	I	V	P	F	F	K	Q	I	A	L	P	L	G	V	V	G	F	L	V	L	S	Y	L	T	I	V	G	T	S	N	A	V	N	L	T	D	G	L	D	G	L	A
	150	160	170	180	190	200																																																						

	40	50	60	70	80	90
m088.pep	TFPVVLVAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVIFCTAMCGACLGFLWFNAYPA					
	:					
g088	AFFPVLVAAGLAIFAYVSGHYQFSQYLQLPYVAGANEVAIFCTAMCGACLGFLWFNAYPA					
	210	220	230	240	250	260
	100	110	120	130	140	150
m088.pep	QVFMDVGALALGAALGTVAIVIRQEFVLVIMGGLFVVEAVSVMLQVGWYKTKKRIFLM					
g088	QVFMDVGALALGAALGTVAIVIRQEFVLVIMGGLFVVEAVSVMLHVGWYKTKKRIFLT					
	270	280	290	300	310	320
	160	170	180	190	200	
m088.pep	APIHHHYEQKGWKETQVVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRHLNAQX					
g088	APIHHHYQLRCWKETQVVVRFWIITIVVVLIGLSTLKIRGNAYVRTPFRRHLNAQX					
	330	340	350	360	370	

```

a088.seq
1   ATGTTTTTAT  GGCTCGCACA  TTTCAGCAAC  TGGTTAACCG  GTCTGAATAT
51  CTTTCAATAC  ACCACATTCC  CGCCGGTCAT  GGCGCGGTTG  ACCCGCTTGG
101 CGTTTTCCCT  GATGTTTCGG  CCGTGCAGTA  TACGCAAGGT  GACCGCGCTG
151 AAATGCGGGC  AGGCAGTGCG  TACCGACGGT  CCGCAAACCC  ACCTCGTCAA
201 AAACGCGACG  CCGACGATGG  GCGGTTTCGT  GATTCTGACC  GCCATTACCG
251 TGTCCACCCT  GTTGTGGGGC  AACTGGGCAA  ACCCGTATAT  CTGGATTCTC
301 TTGGGCGTAT  TGCTGGCCAC  GGGCGCACTC  GGTTTTACG  AGCACTGGCG
351 CAAAGTCGTC  TATAAAGACC  CCAACGGCGT  GTCCGCCAAA  TTCAAATGG
401 TGTGGCAGTC  AAGCGTTGCC  ATATTCGCCG  GTTTTGGCAT  GTTTTACCTT
451 GCGGCCAATT  CCGCCAACAA  TATTTTGATT  GTCCCGTTCT  TCAAACAATT
501 CGCCCTTGCC  CTGGGCGTGG  TCGGCTTTT  GTGTGTGTCT  TACCTGACCA
551 TCGTCGGCAC  ATCCAATGCC  GTCAACCTCA  CCGACGGCTT  GGACGGCCTT
601 GCGACCTTCC  CCGTCGTCTC  CTGTGCCGCC  GGCTTCGCCA  TCTTCGCCTA
651 TGCCAGCGGC  CACTCACAAT  TTGCCCAATA  CCTGCAATTA  CCTTACGTTG
701 CGGGCGCAAA  CGAAGTGGTG  ATTTTCTGTA  CGGCCATGTG  CGCCGCGTGC
751 CTCGGTTTCT  TGTGGTTTAA  CGCCTATCCC  GCGCAAGTCT  TTATGGGCGA
801 TGTGCGTGCA  TTGGCAATTG  GTGCGCGCTC  CGGTACCGTC  GCGCTCATCG
851 TCCGCCAAGA  GTTTGTCTCT  GTCAATTATG  CGGGATTATT  TGTCGTAGAA
901 GCGGTATCCG  TTATGCTTCA  GGTCCGGCTG  TATAAGAAAA  CCAAAAAACG
951 CATCTTCTGT  ATGGCGCCCA  TCCATCACCA  CTACGAACAA  AAAGGCTGGA
1001 AAGAAACCCA  TAGTCGTGTC  CGCTTTTGGA  TTATTACCAT  CGTCTTGGTG
1051 TTGATCGGTT  TGATACCCCT  CAAATCCGC  TGAACCTATG  CCGTCTGAAC
1101 ACCTTTTCAG  CGGCATTTGA  ACGGCAATA  A

```

a088 pep

1	MFLWLAHFSN	WLTGLNIFQY	TFTRAUMAAL	TALAFSLMFG	PWTIRRLTAL
51	KCGQAVRTDG	PQTHLVKNGT	PTMGSSLILT	AITVSTLLWG	NWANPYIWL
101	LGVL L ATGAL	GFYDDWRKVT	YKDPNGVSAS	KFMVWQSSVA	IIAGLALFYL
151	AANSANNILI	VPFFKQIALP	LGVGFLVLVS	YLTIVGTSNA	VNLTDGLDGL
201	<u>ATFPVVLVAA</u>	<u>GLAIFAYASG</u>	HSQFAQYLQL	PYVAGANEVV	IFCTAMCGAC
251	LGFLWFNAYP	AGVFMGDVGA	LALGAALGYV	AVIVRQEFVL	VIMGGLFVVE
301	<u>AVSVMLQVGW</u>	<u>YKKTKKRIFL</u>	MAPIHHHHYE	KGWKETQVVV	RFWIITIVLV
351	LIGLSTLKIR	*TYAV*TPFR	RHLNAQ*		

```

                150      160      170      180      190      200
m088.pep  XXXXXXXXXXXXXXXXXXXXXXXXXXXXGVVGFLVLSYLTIVGTSNAVNLTDLGLDGLA
                |||||
a088      IAGLALFYLAANSANNILIVPFKQIALPLGVVGFLVLSYLTIVGTSNAVNLTDLGLDGLA

```


283

	150	160	170	180	190	200
	210	220	230	240	250	260
m088.pep	TFPVVLAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVFCTAMCGACLGFLWFNAYPA					
a088	TFPVVLAAGLAIFAYASGHSQFAQYLQLPYVAGANEVVFCTAMCGACLGFLWFNAYPA					
	210	220	230	240	250	260
	270	280	290	300	310	320
m088.pep	QVFMGDVGALALGAALGTAVIVRQEFVLVIMGGLFVVEAVSVMQVGWYKTKKRIFLM					
a088	QVFMGDVGALALGAALGTAVIVRQEFVLVIMGGLFVVEAVSVMQVGWYKTKKRIFLM					
	270	280	290	300	310	320
	330	340	350	360	370	
m088.pep	APIHHHYEQGWKETQVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRLHNAQX					
a088	APIHHHYEQGWKETQVVRFWIITIVLVLIGLSTLKIRXTYAVXTSFRRLHNAQX					
	330	340	350	360	370	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 289>:

```

g089.seq
1   ATGCCGCCCA AAATCACGAA GAGCGGGTTT TGCAAACCGG CAATCGCGGC
51  GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATG AATACCACGC
101 CGTTTTTCTC GCCGATTTT TCCACACGGT GCGGCAAGCC TTGGAAGGTT
151 TTGACGTGTT CCAGCAATGC TTCGCGCGGC AAACCGACGG CCTCGCACAA
201 AGCCACGGCA GCCATAACGT TGGCGGCGTT GTGCAAACCT TGCAGCGGGA
251 TGTCTTGCGT AGAAATCAAA TCTTCATTGC CTTGTTTTAA ACAGCCCGTC
301 CCGCGTTCCA ACCAAAAATC GGCTTCGTGT TCCAAGGAAA ACCGTTTCAC
351 TTCACGCCCT GCCCGTTTCA TGGCGCGGCA GAACACGTCG TCCGCATTCA
401 AAACCTGCAC TCCATCGCCA CGGAAAATCT CGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 290; ORF 089.ng>:

```

g089.pep
1   MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFSPIF STRCGKPKWV
51  LTCSSNASRG KPTASHKATA AITLAALCKP CSGMSCVEIK SSLPCFKQPV
101 PRSNQKSASC SKENRFTSRP ARFMARQNTS SAFKTCTPSP RKISALVCA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 291>:

```

m089.seq
1   ATGCCGCCCA AAATCACKAw GAGCGGATTT TGCAAACCGG CAATCGCGGC
51  GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA AACACCACGC
101 CGTTTTTCTC GCCGATTTT TCCACGCGGT GCGGCAGGCC TTGGAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG CCTCACACAA
201 AGCCACkGCA GCCATGACGT TAGCGGCGTT GTGCAKACCT TGCAACGGwA
251 TGTCTTGCGT GACAATCAAA TCTTCATTGC CTTGTTTCAG GCGGCCTGTC
301 TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA ACCATTTTAC
351 CTCGCGCCCC GCACGCTTCA TCGCGCGGCA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC GCCGTCGCCA CGGAAAATCT TGGCTTTGGT ATGCGCATAG

```

This corresponds to the amino acid sequence <SEQ ID 292; ORF 089>:

```

m089.pep
1   MPPKITXSGF CKPAIAAAVA PTFVPLLSSI NTTPFFSPIF STRCGRPWKV
51  LTCSSNASRD KPMASHKATA AMTLAALCXP CNGMSCVTIK SSLPCFRRPV
101 LRSNQSASC SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 089 shows 88.6% identity over a 149 aa overlap with a predicted ORF (ORF 089.ng) from *N. gonorrhoeae*:

m089/g089

284

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPAIAAAVAPTFVPLLSSINTTFFFSPIFSTRCGRPWKVLTCSNASRD					
g089	MPPKITXSGFCKPAIAAAVAPTFVPLLSSMNTTFFFSPIFSTRCGRPWKVLTCSNASRG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m089.pep	KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRNQKSASCSNENHFTSRP					
g089	KPTASHKATAAITLAALCKPCSGMSCVEIKSSLPCFQPVPRSNQKSASCSKENRFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARQNASSAFKTCTPSPRKILALVCAX					
g089	ARFMARQNTSSAFKTCTPSPRKISALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 293>:

a089.seq

```

1  ATGCCGCCTA AAATCACGAA GAGCGGATTT TGCAAACCGG CAATCGCGGC
51  GGCGGTCGCA CCGACGTTTC TGCCTTTGCT GTCGTCGATG AACACCACGC
101 CATTTTTC TC GCGATTTT TCCACGCGGT GCGGCAGGCC TTGAAAGGTT
151 TTGACGTGTT CGAGCAATGC TTCGCGCGGC AAACCGACGG CTTCGCACAA
201 GGCAACGGCA GCCATCACGT TAGTGGCGTT GTGCAAGCCT TGCAGCGGAA
251 TATCTTGGCT GGCAATCAAA TCTTCATTGC CTGTTTCAG GCGACCTGTC
301 TCACGTTCCA ACCAAAAATC GGCTTCGTAT TCCAACGAAA ACCATTTCAC
351 CTCGCGCCCG GCGCGCTTCA TCGCACGACA GAACGCATCG TCCGCATTCA
401 AAACCTGCAC ACCGTCGCCA CGGAAATCT TGGCTTTGGT ATGCGCGTAG

```

This corresponds to the amino acid sequence <SEQ ID 294; ORF 089.a>:

a089.pep

```

1  MPPKITKSGF CKPAIAAAVA PTFVPLLSSM NTTPFFFSPIF STRCGRP*KV
51  LTCSSNASRG KPTASHKATA AITLVALCKP CSGISCVAIK SSLPCFRRPV
101 SRSNQKSASY SNENHFTSRP ARFIARQNAS SAFKTCTPSP RKILALVCA*

```

m089/a089 91.9% identity over a 149 aa overlap

	10	20	30	40	50	60
m089.pep	MPPKITXSGFCKPAIAAAVAPTFVPLLSSINTTFFFSPIFSTRCGRPWKVLTCSNASRD					
a089	MPPKITXSGFCKPAIAAAVAPTFVPLLSSMNTTFFFSPIFSTRCGRPXKVLTCSNASRG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m089.pep	KPMASHKATAAMTLAALCXPCNGMSCVTIKSSLPCFRRPVSRNQKSASCSNENHFTSRP					
a089	KPTASHKATAAITLVALCKPCSGISCVAIKSSLPCFRRPVSRNQKSASYSNENHFTSRP					
	70	80	90	100	110	120
	130	140	150			
m089.pep	ARFIARQNASSAFKTCTPSPRKILALVCAX					
a089	ARFIARQNASSAFKTCTPSPRKILALVCAX					
	130	140	150			

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 295>:

g090.seq

```

1  ATGCCGCTAG TCGAGCAAAT CGTCGTAGCG GTCGAGATGG TCTTCGGAAA

```

q090.pap

```

1  MRVVEQIVVA VEMVFGNVHH RRRSRAQAFG VFQLEAGKLP HPHVRLFAFA
51  LQFCLQDGRT DIARNDIGIQP ALDAEIADQA GYRGFAVAAG NRNHLVAAAV
101 HNVROQLDVA KHAXRRFA*

```

m090.seq

1	ATGCGCATAG	TCGAGCAAGT	CGTCGTAGGG	GTCGAGATGG	TCTTCGGA
51	TGTTACAGCA	CGTCGCGCGA	GTCGGAACGA	GGCTTTCGGT	GTTTTCAGT
101	TGGAAGCTGG	AAAGCTCCAA	CCGCCACAGC	TCGCGCTTTT	TGCTCTCGCG
151	TGCCATTCC	CGACTCCAAA	CCGGCGTGCC	GATATTGCCG	CGGATAACGG
201	TATCCAGCCC	GCACTTGATA	CAGAGATAGC	CGACCAGGCT	CGTTACCGTG
251	GTTTTGCGGT	TGCTGCCGGT	AATCGCAATT	ACCTTGTCGT	CCCGGCCGTT
301	CAACAATGTCC	GCCAGCAATT	CGATGTCGCC	CAACACGCGT	.CCGCCGTTT
351	TGCTGTGA				

m090.pcp

1 MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
51 LPFRLQNRRA DIARDNGIQP ALDTEIADQA RYRGFAVAAG NRRNYLVVPAV
101 HNVRRQFDVA QHAXRRFA*

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 090 shows 83.9% identity over a 118 aa overlap with a predicted ORF (ORF 090.ng) from *N. gonorrhoeae*:

m090/g090

	10	20	30	40	50	60
m090.pep	MRIVEQVVVAVEMVFGNVQHRRRSRTQAAGVFQLEAGKLQHPHVRLFAFALPFRLQNRRRA					
	: : : : :					
g090	MRVVEQIVVAVEMVFGNVHHRSSRAQAAGVFQLEAGKLPHPHVRLFAFALQFCLODGR					
	10	20	30	40	50	60
	70	80	90	100	110	119
m090.pep	DIARDNGIQPALDTEIADQARYRGFAVAAGNRNLYLVPVHNVRQQFDVAQHAXRRFX					
	: : : :					
g090	DIARNDGIQPALDAEIDAQGYRGFAVAAGNRNLVAAAVHNVRQLDVAXHAXRRFX					
	70	80	90	100	110	

a090.seq

1	ATGCGCGTAG	TCGAGCAAGT	CGTCGTAGCG	GTCGAGATGG	TCTTCGGAAA
51	TGTTACAGCAC	TGTCGCGCGCA	GTCGGGCGCA	GGCTTTCGGT	GTTTTCCAGT
101	TGGAAACTGG	AAAGCTCCAA	CACCCACAGC	TCCGCTCTTT	TGCCTTCGCG
151	CTGCAATTCC	GCCTCCAAAA	CCGGCGCGCC	GATATTGCC	CGCATAACGG
201	TATCCAGCCC	ACACTTGATG	CAGAGATAGC	CGACCAGGCT	CGTTACCGTG
251	GTTTTGCCGT	TGCTGCCGGT	AATCGCAATC	ACCTTGTGCG	CGCGGCGGTT
301	CAACAATTGC	GCCAGCAATT	CGATGTGCC	CAACACGCGT	C. CGCCGTTT
351	CGCTTAA				

This corresponds to the amino acid sequence <SEQ ID 300: ORF 090.a>:

286

a090.pep

```

1 MRVVEQVVVA VEMVFGNVQH CRRSRAQAFG VFQLETGKLQ HPHVRLFAFA
51 LQFRLQNRRA DIARDNGIQP TLDAEIQDQA RYRGFAVAAG NRNHLVAAAV
101 HNVRQQFDVA QHAXRRFA*

```

m09/a090 91.5% identity over a 117 aa overlap

	10	20	30	40	50	60
m090.pep	MRIVEQVVAV	EMVFGNVQHR	RRSRTQAFG	VFQLEAGKLQ	HPHVRLFAF	ALPFRLQNRRA
a090	MRVVEQVVVA	VEMVFGNVQH	CRRSRAQAFG	VFQLETGKLQ	HPHVRLFAF	ALQFRLQNRRA
	10	20	30	40	50	60

	70	80	90	100	110	119
m090.pep	DIARDNGIQP	ALDTEIQDQA	RYRGFAVAAG	NRNHLVVP	PAVHNVRQQF	DVAQHAXRRFAX
a090	DIARDNGIQP	TLDAEIQDQA	RYRGFAVAAG	NRNHLVAAV	HNVRQQFDVA	QHAXRRFAX
	70	80	90	100	110	

The following partial DNA sequence was identified in *N. gonorrhoeae*

g090-1.seq This sequence contains multiple stop codons (not shown)

This corresponds to the amino acid sequence < ORF 090-1.ng>:

g090-1.pep (not shown)

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 2>:**m090-1.seq**

```

1 ATGACGGCGT TTGCATTTCA GACGGCATCA CAAAGCCTTA AACGCTTCGA
51 TAAACACTTC CGAACGGTGC GCGTAGCCTT TGAACATATC AAAGCTCGCG
101 CAGGCGGGGC TGAGCAACAC AATATCGCCT GCTTCGGCTT GGGCATATGC
151 CGTCTGAACG GCTTCTCCCA AAGTGGCGCA GTCGGTCATA TTCAAGCCGC
201 AGCCGTCCAA ATCGCGGCGG ATTTGCGGCG CATCGACACC AATCAAGAAC
251 ACGCCTTTTG CTTGCTTAC CAGTGCATCG CGCAGGGGCG TGAAGTCTCG
301 CCCTTTACCC ATGCCGCCCA AAATCAGCAA GAGCGGATTT TGCAAAACCGG
351 CAATCGCGGC GGCAGTCGCG CCGACATTCG TGCCTTTGCT GTCGTCGATA
401 AACACCAACG CGTTTTTCTC GCCGATTTT TCCACGCGGT GCGGCAGGCC
451 TTGGAAGGTT TTGACGTGTT CGAGCAATGC TTCGCGCGAC AAACCGATGG
501 CCTCACACAA AGCCACGGCA GCCATGACGT TAGCGGCGTT GTGCAGACCT
551 TGCAACGGAA TGTCTTGCGT GACAATCAAA TCTTCATGCG CTGTGTTTCAG
601 GCGGCGCTGTC TCGCGTTCCA ACCAGAAATC AGCTTCGTGT TCCAACGAAA
651 ACCATTTTAC CTCGCGCCCG GCACGCTTCA TCGCGCGGCA GAACGCATCG
701 TCCGCATTCA AAACCTGCAC GCCGTGCGCA CGGAAATCT TGGCTTTGGT
751 ATGCGCATAG TCGAGCAAGT CGTCGTAGCG GTCGAGATGG TCTTCGAAA
801 TGTTTCAGCAC CGTCGCCGCA GTCGGACGCA GGCCTTCGGT GTTTTCCAGT
851 TGGAAGCTGG AAAGCTCCAA CACCCACACG TCCGCCTTTT TGCCTTCGCG
901 CTGCCATTCC GCCTCCAAAA CCGGCGTGCC GATATTGCC CCGATAACGG
951 TATCCAGCCC GCACTTGATA CAGAGATAGC CGACCAGGCT CGTTACCGTG
1001 GTTTTGCCGT TGCTGCCCGT AATCGCAATT ACCTGTGCGT CCCGGCGGTT
1051 CACAATGTCC GCCAGCAATT CGATGTGCGC CAACACGCGT CCGCCGTTTT
1101 GCTTGAACGC CTCAATATCC GGCTGCCGCT CGCTGATGCC GGGACTGAGA
1151 GCCAGAATAT CGAAACCGTT GTCCAGCGCA TCTTTCAGAC GGCCCGTGTA
1201 AAACACCAAC CCGTCAAACA TCTTACCGAT TTGCGACACG CGTTCCGGCT
1251 TCAGCTCCGC ATCATACGCA GCAACCTCCG CGCCGTTTTT GCGCAGGTAG
1301 GCAATCATGG AAATACCCGT ACCGCCGAGT CCGCGCAGCA GGATTTTTTT
1351 GTTTTGAAAA GTCATTTTGG TTTGTCCTAA

```

This corresponds to the amino acid sequence <SEQ ID 3; ORF 090-1>:

m090-1.pep

```

1 MTAFAFQTAS QSLKRFDKHF RTVRVAFEH I KARAGGAEQH NIACFGLGIC
51 RLNGFSQSCA VGHIAAAVQ IAADLRRIDT NQEHAFCLAY QCIAQGREV L
101 PETHAAQNHE ERILQTGNRG GSRADIRAF VVDKHHAVFL ADFFHAVRQA
151 LEGFVDFEQC FARQTDGLTQ SHGSHDVSGV VQTLQRNVLR DNQIFIALFQ
201 AACLAEPQEI SFVFQRKPFY LAPGTLHRAA ERIVRIQNLH AVATENLGFG
251 MRIVEQVVVA VEMVFGNVQH RRRSRTQAFG VFQLEAGKLQ HPHVRLFAFA
301 LPFRLQNRRA DIARDNGIQP ALDTEIQDQA RYRGFAVAAG NRNHLVVP
351 HNVRQQFDVA QHASAVLLER LNIRLPLADA GTESQNIETV VQRIQTARV
401 KHQPVKHLTD LRHAFRLQLR IIRSNLRAVF AQVGNHGNTR TAESGDEDF
451 VLKSHFGLS*

```

This corresponds to the amino acid sequence <SEQ ID 304; ORF 091.ng>:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 305>:

This corresponds to the amino acid sequence <SEQ ID 306; ORF 091>:

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 091 shows 84.2% identity over a 101 aa overlap with a predicted ORF (ORF 091.ng) from *N. gonorrhoeae*:

m091/g091

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 307>:

```

a091.seq
1  ATGGAATAC  CCGTGCCGCC  AAGTCCGGCG  ACAGGATT  TTTGGTTTG
51  GAAATCATT  TGGTTTGTC  TAAACAAAT  CATATTGAC  AGGGGATGC
101 TGATCTGCT  CAAGCCGCT  TCAGACGCA  TCGGGAGCT  TTCAATAACC
151 CGCTTTCAG  CGTTGGTCAT  TGTGCGAGCT  GTCTTGCTAT  CCGTTTGGAC
201 AAGCCATTGCC  AAGCCATTCT  TGTGCAAGGG  CGCGGTCTTG  GCGCACGCCG
251 CGTCTTTCGG  CATACTATC  GGCCAAATTG  CTTTGGGG

```

This corresponds to the amino acid sequence <SEQ ID 308; ORF 091.a>:

a091.pep

```

1  MEIPVPPSPA TRIFLFWKSF WFVLKQIILS RGCLILLKPL SDGIASCSIT
51 RFQALVIVAA VLVSVLTSLA KPFLCKGAVL AHAASFGIHH AQIVLG

```

m091/a091 93.8% identity over a 96 aa overlap

```

              10      20      30      40      50      60
m091.pep  MEIPVPPSPATRIFLFEKSFWFVLKQIILSRRCPPPLPKPLSDGIASCSITRLQALVIVAA
          |||||
a091      MEIPVPPSPATRIFLFWKSFWFVLKQIILSRGCLILLKPLSDGIASCSITRFQALVIVAA
              10      20      30      40      50      60

              70      80      90      100
m091.pep  VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLGLGYPLR
          |||||
a091      VLVSVLTSLAKPFLCKGAVLAHAASFGIHHAQIVLG
              70      80      90

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 309>:

g092.seq

```

1  ATGTTTTTTA TTTCAATCCG CTATATATTT GTCAGAAAAC TATGGTGCGC
51  AAACGGTCAG ACCTTTAAAA TAACGCCTTT ACGCACTAAA AACCAACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTAA GCAACATCCA TTTTGTGCGT
151 ATCGGCGGCG TCGGCATGAG CGGTATCGCC GAAGTCTTGC ACAATTTGGG
201 CTTTAAAGTT TCCGGTTCGG ATCAGGCGCG AAATGCCGCT ACCGAGCATT
251 TGAGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC AGAACACGTT
301 AACGGTgcgg ATGTCGTCGT TGCCTCTACC GCCGTCAAGA AAGAAaatcC
351 CGAAGTgtgc gctGCGTTGG AGCGGCAAT TCCCGTTATT CCGCGCGCCT
401 TGATGCTGGC AGAGCTGATG CGCTTCCGTG ACGgcatcgc cattgccggT
451 ACGCACGGCA AAACCAACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTC GACCCCACTT TCGTTATCGG CGGCAAACTC AACGCCGAG
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAATCCGATG CCTCTTTCCT ACATCTGACC CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTGTTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAAACCT TATGCTACTT ACGGTTTGGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GGCGCGCAAA TGAAATTACG CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTGGC AGCCATCGGc gtggcGCTgg
1001 aagtcGgCGC ATcggttgAA GCGAtcCAAA AaggCTTGCT CGGCTTTGAA
1051 GGGCTCGGCC GCCGCTTCCA AAAATAcggc gacatCAagt tgccaaacgg
1101 cggGaccgCT TTgctGGTGG ACGATTAcgg ACACCACCCC GTCGAAATGG
1151 CGGcaaccct tgccgcTGCA CGCGGCGCGT ATCCGAAAAA acgtTTGGTG
1201 CtcgCCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTGAAGA
1251 CTTTACCAAA GACTCAATA CCGTTGatgC GCTGGTACTG ACCGAAGTTT
1301 AtgccgcgG CGAAGAGCCG GTTGCCCGCG CCGactcCCG CGCCTTGGCG
1351 CGTGCTATCC GCGTATTGGG CAAACTTGAG CCGATTACTT GCGAAAatgt
1401 cgccgACCTG CCGCAAATGC TGATGAATGT TTTACAGGAT Gcggatgtg
1451 tgttgAATAT GggTgcggga agcatcaacc gcgttccttc cgcgctgttg
1501 gaattgtcga AACAGatttg A

```

This corresponds to the amino acid sequence <SEQ ID 310; ORF 092.ng>:

g092.pep

```

1  MFFISIRYIF VRKLWCANGQ TFKITPLRTK NQPERNIMMK NRVSNIHFGV
51  IGGVGMSGIA EVLHNLGFKV SGSDQARNAA TEHLSSLGIQ VYPGHTAEHV
101 NGADV VVAST AVKKENPEVV AALERQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK

```

289

```

251 AFLCVDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEV GASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP VAAADSRALA
451 RAIRVLGKLE PIYCENVADL PQMLMNVLQD GDVVLNMGAG SINRVPSALL
501 ELSKQI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 311>:

m092.beq

```

1 ATGTTTTTTA TTCAATCCG CTATATATTT GTCAGAAAAC TATGGCGCGC
51 AACCGGTGAG CCCTTTAAAA TAACGCCTTT ACGCATCGAA AATCCACCGG
101 AACGCAACAT TATGATGAAA AATCGAGTTA CCAACATCCA TTTTGTCTGGT
151 ATCGGCGGCG TCGGCATGAG CGGCATCGCC GAAGTCTTGC ACAATTGGG
201 CTTTAAAGTT TCCGGTTCGG ATCAgGCGCG AAATGCCGCT ACCGAGCATT
251 TGGGCAGCCT GGGCATTCAA GTTTATCCCG GCCATACCGC CGAACACGTT
301 AACGGTGGCG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCCGTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGCTTCCGTG ACGGCATCGC CATTGCCGGC
451 ACGCACGGCA AAACACGAC CACGAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCGC
551 GCACTAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGC GTCGAAAAAC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTATTGACAG CGAACACGTC CGCGCGATT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGGG CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCAAAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTCGTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGCGCG ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GCGCTCGGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 CGGGACCGCG CTCTTGGTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CGGCGACCCCT TGCCGCCGCA CGCGGCGCGT ATCTGAAAA ACGTTTGGTA
1201 CTCGCCCTTC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCCGG TGAAGAGCCG ATTGCCGCCG CCGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTGGG CAAACTCGAG CCGATTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC GTTTGAACGT TTTGCAGGAC GGCGACATCG
1451 TGTTGAATAT GGGCGCGGGA AGCATCAACC GCGTCCCCGC CGCGCTGCTG
1501 GCATTGTGCA AACAGATTG A

```

This corresponds to the amino acid sequence <SEQ ID 312; ORF 092>:

m092.pep

```

1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNIHFVG
51 IGGVGMMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101 NGADVVTST AVKKENPEV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEV GASVE AIQKGLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLAAA RGAYLEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVADL PEMLLNVLQD GDIVLNMAG SINRVPAALL
501 ALSKQI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 092 shows 96.6% identity over a 506 aa overlap with a predicted ORF (ORF 092.ng) from *N. gonorrhoeae*:

m092/g092

```

10      20      30      40      50      60
m092.pep  MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMMSGIA

```

a092.seq

1	ATGTTTTTTA	TTTCAATCCG	CTATATATTT	GTCAGAAAAC	TATGGCGCGC
51	AAACCGCTCAG	CCCTTTAAAA	TAACGCGCTT	ACGCATCGAA	AATCCACCGG
101	AACGCAACAT	TATGATGAAA	ACGTGAGTA	CCAAATACCA	TCTTTCGGTT
151	ATCGGCGGCG	TCGGCATGAG	CGGTATCGCC	GAAGTCTTGC	ACAATTGTGG
201	TTTTAAAGTT	TCGGGTTCCG	ATCAGGCGCG	AAATGCCGCT	ACCGAGCATT
251	TGGGCAACCT	GGGCATTCCA	GTTTATCCCG	GCCATACGCA	AGCAACAGTT


```

301 AACGGTGC GG ATGTCGTCGT TACCTCTACC GCCGTCAAAA AAGAAAATCC
351 CGAAGTTGTC GCTGCGTTGG AGCAGCAAAT TCCC GTTATT CCGCGCGCCC
401 TGATGTTGGC GGAGTTGATG CGTTCCGTG ACGGCATCGC CATTGCCGGC
451 ACGCACGGCA AAACCACGAC CACCAGCCTG ACCGCCTCCA TCCTCGGCGC
501 GGCAGGACTT GACCCGACTT TCGTTATCGG CGGCAAACTC AACGCCGCAG
551 GCACCAACGC CCGCTTGGGC AAAGGCGAAT ACATCGTTGC CGAAGCCGAC
601 GAGTCGGATG CATCCTTTCT GCACCTGACA CCGATTATGT CCGTCGTTAC
651 CAATATCGAC GAAGACCATA TGGATACCTA CGGGCACAGT GTTGAGAAGC
701 TGCATCAGGC GTTTATCGAT TTCATCCACC GTATGCCCTT CTACGGCAAA
751 GCCTTTTGTG GTATTGACAG CGAACACGTC CGCGCGATTT TGCCCAAAGT
801 GAGCAAACCT TATGCTACTT ACGGTTTGA CGATACCGCC GACATCTACG
851 CCACCGACAT CGAAAACGTC GCGCGCAAAA TGAAATTCAC CGTCCATGTT
901 CAAATGAAAG GACATGAGCA GGGGTGCTTT GAAGTCGTGC TGAATATGCC
951 CGGCAGACAC AACGTGCTGA ACGCATTTGGC AGCCATCGGC GTGGCGCTGG
1001 AAGTCGGCGC ATCGGTTGAA GCGATCCAAA AAGGCTTGCT CGGCTTTGAA
1051 GGTGTGCGCC GCCGCTTCCA AAAATACGGC GACATCAAGT TGCCAAACGG
1101 TGAACCGCGC CTCTTGCTGG ACGACTACGG ACACCACCCC GTCGAAATGG
1151 CCGCGACCCCT TTCCGCGCGC ACGCGCGCGT ATCCGGAAAA ACCTTTGGTA
1201 CTCGCTTCC AGCCGCACCG CTATACCCGC ACGCGCGATT TGTTTGAAGA
1251 CTTTACCAAA GTCCTCAATA CCGTTGACGC GCTGGTGCTG ACCGAAGTTT
1301 ATGCCGCGCG TGAAGAGCCG ATTGCCGCGC CTGATTCCCG CGCTCTTGCC
1351 CGCGCCATCC GCGTGTGGG CAAACTCGAG CCGATTACT GCGAAAACGT
1401 TGCCGATCTG CCCGAAATGC TGTGTAACGT TTTGCAGGAC GGCGACATCG
1451 TGTGTAATAT GGGTGCGGGA AGCATCAACC GCGTCCCGC CGCGCTGCTG
1501 GAATTGTCGA AACAGATTG A

```

This corresponds to the amino acid sequence <SEQ ID 314; ORF 092.a>:

a092.pep

```

1 MFFISIRYIF VRKLWRANGQ PFKITPLRIE NPPERNIMMK NRVTNHIFVG
51 IGGVGMMSGIA EVLHNLGFKV SGSDQARNAA TEHLGSLGIQ VYPGHTAEHV
101 NGADVVTST AVKKENPEVV AALEQQIPVI PRALMLAELM RFRDGIAG
151 THGKTTTSL TASILGAAGL DPTFVIGGKL NAAGTNARLG KGEYIVAEAD
201 ESDASFLHLT PIMSVVTNID EDHMDTYGHS VEKLHQAFID FIHRMPFYGK
251 AFLCIDSEHV RAILPKVSKP YATYGLDDTA DIYATDIENV GAQMKFTVHV
301 QMKGHEQGSF EVVLNMPGRH NVLNALAAIG VALEVGASVE AIQKLLGFE
351 GVGRRFQKYG DIKLPNGGTA LLVDDYGHHP VEMAATLSAA RGAYPEKRLV
401 LAFQPHRYTR TRDLFEDFTK VLNTVDALVL TEVYAAGEEP IAAADSRALA
451 RAIRVLGKLE PIYCENVDL PEMLLNVLQD GDIVLNMAG SINRVPALL
501 ELSKQI*

```

m092/a092 99.4% identity over a 506 aa overlap

	10	20	30	40	50	60
m092.pep	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMMSGIA					
a092	MFFISIRYIFVRKLWRANGQPFKITPLRIENPPERNIMMKNRVTNIHFVGIGGVGMMSGIA					
	10	20	30	40	50	60
m092.pep	EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNAGADVVTSTAVKKENPEVV					
a092	EVLHNLGFKVSGSDQARNAATEHLGSLGIQVYPGHTAEHVNAGADVVTSTAVKKENPEVV					
	70	80	90	100	110	120
m092.pep	AALQQIPVIPRALMLAELMRFRDGIAGTHGKTTTSLTASILGAAGLDPTFVIGGKL					
a092	AALQQIPVIPRALMLAELMRFRDGIAGTHGKTTTSLTASILGAAGLDPTFVIGGKL					
	130	140	150	160	170	180
m092.pep	NAAGTNARLGKGEYIVAEADESDASFLHLPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					
a092	NAAGTNARLGKGEYIVAEADESDASFLHLPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					
	190	200	210	220	230	240
m092.pep	NAAGTNARLGKGEYIVAEADESDASFLHLPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					
a092	NAAGTNARLGKGEYIVAEADESDASFLHLPIMSVVTNIDEDHMDTYGHSVEKLHQAFID					

292

	190	200	210	220	230	240
	250	260	270	280	290	300
m092.pep	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFVHV					
a092	FIHRMPFYGKAFLCIDSEHVRAILPKVSKPYATYGLDDTADIYATDIENVGAQMKFVHV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m092.pep	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKGLLGFEVGRRFQKYG					
a092	QMKGHEQGSFEVVLNMPGRHNVNLALAAIGVALEV GASVEAIQKGLLGFEVGRRFQKYG					
	310	320	330	340	350	360
	370	380	390	400	410	420
m092.pep	DIKLPNGGTALLVDDYGHHPVEMAATLAAARGAYLEKRLVLA FQPHRYTRTRDLFEDETK					
a092	DIKLPNGGTALLVDDYGHHPVEMAATLSAARGAYPEKRLVLA FQPHRYTRTRDLFEDETK					
	370	380	390	400	410	420
	430	440	450	460	470	480
m092.pep	VLNTVDALVLT EYVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
a092	VLNTVDALVLT EYVYAAGEEPIAAADSRALARAIRVLGKLEPIYCENVADLPEMLLNVLQD					
	430	440	450	460	470	480
	490	500				
m092.pep	GDIVLNMGAGSINRVPAALLALSKQIX					
a092	GDIVLNMGAGSINRVPAALLELSKQIX					
	490	500				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 315>:

```

g093.seq
1  aTGCAAAAtt ttgGCAAAGT ggccgtATTG ATGGGtggtT TTTCCAGCGA
51  ACGAGAAatc tcgcTGGACA GCgGTACCGC CATTtTGAAC GCCTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGACC CTAAGGAAAC GCCGTtATCC
151 GAACTGAAGG AGCGGGGCTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAAGTGT TTGGGCATTC
251 CCTATACCGG CAGCGGTGTC GCCGCCTCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTtGGCA GGCATTGGGA TTACCCGTTC CCGAGTTCGC
351 CGTACTGTAC GATGATACCG ATTTCGATGC CGTCGAAGAA AAATTGGGTC
401 TGCCGATGTT TGTGAAGCCG GCGGCCGAAG GCAGCAGCgt cggcgtggta
451 aAAGTCAAAG AAAaaggccg TCTGAAAAGC GTTtacgaag aatTGAaaCA
501 CCTTcagggg cgaAatcatt gccgAacgTT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATA ACATCATCCC
601 CGCAACCGAG TTTTACGAct acgaagccaa GtaciaCCGA GACGAcacca
651 tttatCAAATG TCCTTCGGAA GATTtGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTCGCGG CGCACAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801 TCAACACCCT GCCCGGTATG ACCGGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 316; ORF 093.ng>:

```

g093.pep
1  MQNFGKQAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKERGFQTA FNILHGTYGE DGA VQGALEL LGIPYTGSGV AASAIMGDKY
101 RCKLIWQALG LPVPEFAVLY DDTDFDAVEE KLGLPMFVKP AAE GSSVGVV
151 KVKEKGR LKS VYEELKHLQG R NHCRTFYRR RRIFLPRPER QRAARHTHLP
201 RNRVLR LRSQ VQPRRHLSM SFGRFDRSRR KPDARTGGS RTGNRCGR LR
251 ARRFPPQRYRR QTL SVGNQHP ARYDRP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 317>:

m093.seq

293

```

1  ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CTAAAGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCAA GACGGGGCGG TTCAGGGTGC ATTGGAAGTG TTGGGCATTC
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAC TGATTTGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTCGATGC CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATTGAAACA
501 CCTTCAGGG. CGAAATCATT GCCGAACGTT TTATCGGCGG CGGCGAATAT
551 TCCTGCCCCG TCCTGAACGG CAAAGGGCTG CCCGGCATAC ACATCATTCC
601 CGCAACCGAG TTTTACGACT ACGAAGCCAA GTACAACCGC GACGACACCA
651 TTTATCAATG TCCTTCGGAA GATTTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTTCCGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801 TCAACACCCT GCCCGGTATG ACGAGCCATA G

```

This corresponds to the amino acid sequence <SEQ ID 318; ORF 093>:

```

m093.pep
1  MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGX DGAVQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLGLPMFVKP AEGSSVGVV
151 KVKGKGRLKS VYEELKHLQX RNHCRTFYRR RRIFLPRPER QRAARHTHS
201 RNRVLRLRSQ VQPRRHLSM SFGRFRDSRR KPDARTGGSR RAGNRCGRLR
251 ARRFPQRYRR QTLSVGNQHP ARYDEP*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 093 shows 96.7% identity over a 276 aa overlap with a predicted ORF (ORF 093.ng) from *N. gonorrhoeae*:

m093/g093

	10	20	30	40	50	60
m093.pep	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS ELKAQGFQTA					
g093	MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS ELKERGFQTA					
	10	20	30	40	50	60
m093.pep	70	80	90	100	110	120
	FNILHGTYGX DGA V QGALELLG IPYTGSGVAASAIGMDKY RCKLIWQALGLPVPEFAVLH					
g093	FNILHGTYGEDGA V QGALELLG IPYTGSGVAASAIGMDKY RCKLIWQALGLPVPEFAVL Y					
	70	80	90	100	110	120
m093.pep	130	140	150	160	170	180
	DDTDFDAVEE KLGLPMFVKPAAEGSSVGVV KVKGKGR L KS VYEELKHLQX RNHCR T FYRR					
g093	DDTDFDAVEE KLGLPMFVKPAAEGSSVGVV K E KGR L KS VYEELKHLQGRNHCR T FYRR					
	130	140	150	160	170	180
m093.pep	190	200	210	220	230	240
	RRI F LP R PERQRAARH T HS RNRVLR L RSQVQPRRH L SM SFG R FR D SR R KPDARTGGS R					
g093	RRI F LP R PERQRAARH T HPNRVLR L RSQVQPRRH L SM SFG R FR D SR R KPDARTGGS R					
	190	200	210	220	230	240
m093.pep	250	260	270			
	RAGNRCGR L RRARFPQRYRRQTL S VGNQHPARYDEPX					
g093	RTGNRCGR L RRARFPQRYRRQTL S VGNQHPARYDRPX					
	250	260	270			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 319>:

a093.seq

```
1   ATGCAGAATT TTGGCAAAGT GGCCGTATTG ATGGGCGGTT TTTCCAGCGA
51  ACGAGAAATC TCGCTGGACA GCGGCACCGC CATTTTGAAT GCTTTAAAAA
101 GCAAAGGCAT AGACGCATAC GCCTTCGATC CCAAGGAAAC CCCATTGTCT
151 GAATTGAAGG CACAAGGTTT TCAGACGGCA TTCAACATCC TTCACGGTAC
201 TTACGGCGAA GACGGGGCTG TTCAGGGTGC ATTGGAAGTG TTGGGCATTC
251 CCTATACCGG CAGCGGTGTC GCCGCATCCG CCATCGGCAT GGACAAATAC
301 CGCTGCAAAAC TGATTGGGCA GGCATTGGGA TTGCCCGTTC CCGAGTTCGC
351 CGTCCTGCAC GACGACACTG ATTTTCGATG CGTCGAAGAA AAATTGGGCC
401 TGCCGATGTT TGTGAAACCG GCGGCCGAAG GCAGCAGCGT AGGCGTGGTA
451 AAAGTCAAAG GAAAAGGCCG TCTGAAAAGC GTTTACGAAG AATGAAACA
501 CTTTCAGGG. CGAATCATT GCCGAACGGT TTATCGGCGG CCGCGAATAT
551 TCCTGCCCTG TGTGAACGG CAAAGGCCTG CCCGGCATAC ACATCATCCC
601 CGCGACCGAG TTTTATGACT ACGAAGCCAA GTACAACCGC AACGACACCA
651 TTTATCAATG TCCTTCGGAA GATCTGACCG AAGCCGAAGA AAGCCTGATG
701 CGCGAACTGG CGGTCGCGG CGCGCAGGCA ATCGGTGCGG AAGGCTGCGT
751 GCGCGTCGAT TTCCTCAAAG ATACCGACGG CAAACTCTAT CTGTTGGAAA
801 TCAACACCTT GCCCGGTATG ACCGGCCATA G
```

This corresponds to the amino acid sequence <SEQ ID 320; ORF 093.a>:

a093.pep

```
1   MQNFGKVAVL MGGFSSEREI SLDSGTAILN ALKSKGIDAY AFDPKETPLS
51  ELKAQGFQTA FNILHGTYGE DGA VQGALEL LGIPYTGSGV AASAIGMDKY
101 RCKLIWQALG LPVPEFAVLH DDTDFDAVEE KLG LPMFVKP AAEGSSVG VV
151 KVKGKGR LKS VYEELKHFQX RNHCRTVYRR RRIFLPCVER QRPARTHHP
201 RDRVL*LR SQPQRHHL SM SFG RSDRSRR KPDARTGGS RAGNRCGR LR
251 ARRF PQRYRR QTL SVGNQH PARYDRP*
```

m093/a093 95.7% identity over a 276 aa overlap

	10	20	30	40	50	60
m093.pep	MQNFGKVAVL	MGGFSSEREI	SLDSGTAILN	ALKSKGIDAY	AFDPKETPLS	ELKAQGFQTA
a093	MQNFGKVAVL	MGGFSSEREI	SLDSGTAILN	ALKSKGIDAY	AFDPKETPLS	ELKAQGFQTA
	70	80	90	100	110	120
m093.pep	FNILHGTYGX	DGA VQGALELL	GIPYTGSGVA	ASAIGMDKYR	CKLIWQALGL	VPPEFAVLH
a093	FNILHGTYGED	GAVQGALELL	GIPYTGSGVA	ASAIGMDKYR	CKLIWQALGL	VPPEFAVLH
	130	140	150	160	170	180
m093.pep	DDTDFDAVEE	KLG LPMFVKP	AAEGSSVG VV	KVKGKGR LKS	VYEELKHLQX	RNHCRTFYRR
a093	DDTDFDAVEE	KLG LPMFVKP	AAEGSSVG VV	KVKGKGR LKS	VYEELKHFQX	RNHCRTVYRR
	190	200	210	220	230	240
m093.pep	RRIFLPRPER	QRAARHTH	SRNRLRLRS	QVQPRRHLS	MSFGRFDRS	RKPDARTGGS
a093	RRIFLPCVER	QRPARTHHP	RDRVLXLR	SQVQPRRHLS	MSFGRSDRS	RKPDARTGGS
	250	260	270			
m093.pep	RAGNRCGR LR	ARRFPQRYRR	QTL SVGNQH	PARYDEPX		
a093	RAGNRCGR LR	ARRFPQRYRR	QTL SVGNQH	PARYDRPX		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 321>:

```
g094.seq
  1 ATGTATTTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
  51 GCCGCCGATA ACGAAAGTGG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
 101 CCGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTAcgggc aatggcgatg
 151 cCGTCACGGA AGCGCATCAG CTCTGCCAGC ATCAAGGCGC GCGGAATAAC
 201 GCGAATTTGC CGCTCCAACG CAgcgacaAC TTCGGgattt TCTTTCTTGA
 251 CGGCGGTAGA GGCAACGACG ACATccgcAC CGTTAACGTG TTCTGCGGTA
 301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 322; ORF 094.ng>:

```
g094.pep
  1 MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
  51 PSRKRISSAS IKARGITGIC RSNAATTSGF SFLTAVEATT TSAPLTCSAV
 101 WPG*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 323>:

```
m094.seq
  1 ATGTATTTCGC CTTTGCCCAA GCGGGCGTTA GTGCCTGCGG CGTTGAGTTT
  51 GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
 101 CCGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
 151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
 201 GCGAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATT TCTTTTGA
 251 CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCGGCGGTA
 301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 324; ORF 094>:

```
m094.pep
  1 MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
  51 PSRKINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
 101 WPG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 094 shows 95.1% identity over a 103 aa overlap with a predicted ORF (ORF 094.ng) from *N. gonorrhoeae*:

```
m094/g094

      10      20      30      40      50      60
m094.pep MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAN
|||
g094     MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRISSAS
      10      20      30      40      50      60

      70      80      90     100
m094.pep IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
|:|
g094     IKARGITGICRSNAATTSGFSFLTAVEATTTSAPLTCSAVWPGX
      70      80      90     100
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 325>:

```
a094.seq
  1 ATGTATTTCGC CTTTGCCCAA GCGGGCGTTG GTGCCTGCGG CGTTGAGTTT
  51 GCCGCCGATA ACGAAAGTCG GGTCAAGTCC TGCCGCGCCG AGGATGGAGG
 101 CCGTCAGGCT GGTGGTCGTG GTTTTGCCGT GCGTGCCGGC AATGGCGATG
 151 CCGTCACGGA AGCGCATCAA CTCCGCCAAC ATCAGGGCGC GCGGAATAAC
 201 GCGAATTTGC TGCTCCAACG CAGCGACAAC TTCGGGATT TCTTTTGA
 251 CGGCGGTAGA GGTAACGACG ACATCCGCAC CGTTAACGTG TTCTGCGGTA
 301 TGGCCGGGAT AA
```

This corresponds to the amino acid sequence <SEQ ID 326; ORF 094.a>:

297

```

g095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           10      20      30      40      50      60

           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
g095      HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKEASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
g095      CLRRX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 331>:

```

a095.seq
1  ATGTCCTTTC ATTGAACAT GGACGGTGAA TTTCATTGTC GCGCCGACGT
51 TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAATC GCGCGGACGT GTTCGCTGTC
151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGCTTC TCAACACTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301 GATGCATCG ACCGTCGGCT TCGGCAACGA TGTATTGCC TTTGCCCAAG
351 CGGGCGTTGG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 332; ORF 095.a>:

```

a095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51 NTQKGFAVEG HTVDEIDKRL MQLLNTVPVG IHMFVDIGN DGHNRCQCRK
101 DASDRRLRQR CIRLCPSGRW CLRR*

```

m095/a095 96.0% identity in 124 aa overlap

```

           10      20      30      40      50      60
m095.pep  MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           |||||
a095      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
           10      20      30      40      50      60

           70      80      90      100     110     120
m095.pep  HTVDEIDKRLMQFFDAVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRX
           |||||
a095      HTVDEIDKRLMQLLNTVPVGIHMFVDIGNDGHNRCQCRKDASDRRLRQRCIRLCPSGRW
           70      80      90      100     110     120

m095.pep  CLRRX
           ||||
a095      CLRRX

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 333>:

```

g096.seq
1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51 CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGTGCGC CGCAACCGC CAGTTCGCGC ATCAGGCTTT CTTCCGGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGAtaaat ggtgTCGTCT CGGttgtaCt

```

a094.pep

```

1  MYSPLPKRAL VPAALSLPPI TKVGSSPAAP RMEAVRLVVV VLPCVPAMAM
51  PSRKRINSAN IRARGITGIC CSNAATTSGF SFLTAVEVTT TSAPLTCSAV
101 WPG*

```

m094/a094 100.0% identity over a 103 aa overlap

```

              10      20      30      40      50      60
m094.pep      MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
               |||||
a094           MYSPLPKRALVPAALSLPPI TKVGSSPAAPRMEAVRLVVVVLPCVPAMAMPSRKRINSAN
              10      20      30      40      50      60

              70      80      90     100
m094.pep      IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
               |||||
a094           IRARGITGICCSNAATTSGFSFLTAVEVTTTSAPLTCSAVWPGX
              70      80      90     100

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 327>:

```

g095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AACACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGGGTCA GTGTAGGAAA
301 GAGGCATCGG ATCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351 CGGGCGTTGG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 328; ORF 095.ng>:

```

g095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRGQCRK
101 EASDRRLRQR CIRLCPGRW CLRR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 329>:

```

m095.seq
1  ATGTCCTTTC ATTTGAACAT GGACGGTGAA TTTCATTTGC GCGCCGACGT
51  TTTCGATGTC GGTGGCGTAG ATGTCGGCGG TATCGTCCAA ACCGTAAGTA
101 GCATAAGGTT TGCTCACTTT GGGCAAAATC GCGCGGACGT GTTCGCTGTC
151 AATACACAAA AAGGCTTTGC CGTAGAAGGG CATACGGTGG ATGAAATCGA
201 TAAACGCCTG ATGCAGTTT TCGACGCTGT GCCCGTAGGT ATCCATATGG
251 TCTTCGTCGA TATTGGTAAC GACGGACATA ATCGGTGTCA GTGCAGAAAG
301 GATGCATCCG ACCGTCGGCT TCGGCAACGA TGTATTCGCC TTTGCCCAAG
351 CGGGCGTTAG TGCCTGCGGC GTTGA

```

This corresponds to the amino acid sequence <SEQ ID 330; ORF 095>:

```

m095.pep
1  MSFHLNMDGE FHLRADVFDV GGVDVGGIVQ TVSSIRFAHF GQNRADVFAV
51  NTQKGFAVEG HTVDEIDKRL MQFFDAVPVG IHMVFDIGN DGHNRGQCRK
101 DASDRRLRQR CIRLCPGRX CLRR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 095 shows 97.6% identity over a 124 aa overlap with a predicted ORF (ORF 095.ng) from *N. gonorrhoeae*:

m095/g095

```

              10      20      30      40      50      60
m095.pep      MSFHLNMDGEFHLRADVFDVGGVDVGGIVQTVSSIRFAHFGQNRADVFAVNTQKGFAVEG
               |||||

```

298

```

201  tggcttcgta gTCGTAAAC TCGGTTGCGG GGATGATGTG TATGCCGGGC
251  AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301  AcgtTcggca atgaTTtcgc ccctgAAGGT GttTCAattc ttcgtaAACG
351  CTTTTCAGAC ggccttTTTC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 334; ORF 096.ng>:

```

g096.pep
1  MAGHTGQGVDFQQIEFAVGI FEEIDAHAAF RTDCLCAANR QFAHQAFFGF
51  GQIFRRTLIN GVVSVVLGFV VVKLGCGDDV YAGQPFQVQD GAGIFAAADK
101 TFGNDFAPEG VSILRKRFSD GLFL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 335>:

```

m096.seq
1  ATGGCTCGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGCGCGC CGCGAACC GCAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAAATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTCG CGGTTGTAAT
201 TGGCTTCGTA GTCGTAAAC TCGGTTGCGG GAATGATGTG TATGCCGGGC
251 AGCCCTTTGC CGTTCAGGAC GGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 ACGTTTCGCA ATGATTTCGC CC.TGAAGGT GTTTCATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 336; ORF 096>:

```

m096.pep
1  MARHTGQGVDFQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF
51  GQIFRRTLIN GVVAVVLGFV VVKLGCGNDV YAGQPFQVQD GAGIFAAADK
101 TFGNDFAXEG VSILRKRFSD GLFL*

```

m096/g096 96.0% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLRAANRQFAHQAFFGFGQIFRRTLIN					
g096	MAGHTGQGVDFQQIEFAVGIFEEIDAHAAFRTDCLCAANRQFAHQAFFGFGQIFRRTLIN					
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDVYAGQPFQVQDGAGIFAAADKTFGNDFAXEGVSILRKRFSD					
	:					
g096	GVVSVVLGFVVVKLGCGDDVYAGQPFQVQDGAGIFAAADKTFGNDFAPEGVSILRKRFSD					
	70	80	90	100	110	120
m096.pep	GLFLX					
g096	GLFLX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 337>:

```

a096.seq
1  ATGGCCGGTC ATACCGGGCA GGGTGTGAT TTCCAACAGA TAGAGTTTGC
51  CGTCGGTATC TTTGAGGAAA TCGACGCGCA CGCAGCCTTC CGCACCATT
101 GCCTGCGCGC CGCGAACC GCAGTTCGCGC ATCAGGCTTT CTTCGGCTTC
151 GGTCAAGATCT TCCGAAGGAC ATTGATAAAT GGTGTCGTTG CGGTTGTAAT
201 TGGCTTCGTA GTCATAAAC TCGGTCGCGG GGATGATGTG TATGCCGGGC
251 AGGCCTTTGC CGTTCACAC AGGGCAGGAA TATTCGCCGC CGCCGATAAA
301 CCGTTCGGCA ATGATTTCGC CCT.GAAAGT GTTTCATTC TTCGTAAACG
351 CTTTTCAGAC GGCCTTTTCC TTTGA

```

This corresponds to the amino acid sequence <SEQ ID 338; ORF 096.ng>:

```

a096.pep
1  MAGHTGQGVDFQQIEFAVGI FEEIDAHAAF RTDCLRAANR QFAHQAFFGF

```


51 GQIFRRTLIN GVVAVVLGFV VIKLGRGDDV YAGQAFVQH RAGIFAAADK
 101 PFGNDFAXES VSILRKRFSD GLFL*

m096/a096 92.7% identity in 124 aa overlap

	10	20	30	40	50	60
m096.pep	MARHTGQGVDFQQIEFAVGIFEEIDAHA	AERTDCLRAANRQFAHQAFFGFGQIFRRTLIN				
a096	MAGHTGQGVDFQQIEFAVGIFEEIDAHA	AERTDCLRAANRQFAHQAFFGFGQIFRRTLIN				
	10	20	30	40	50	60
	70	80	90	100	110	120
m096.pep	GVVAVVLGFVVVKLGCGNDVYAGQPF	FAVQDGAGIFAAADKTFGNDFA	XEGVSILRKRFSD			
a096	GVVAVVLGFVVVKLGRGDDVYAGQAF	VQHRAGIFAAADKPF	GNDFA	XESVSILRKRFSD		
	70	80	90	100	110	120
m096.pep	GLFLX					
a096	GLFLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 339>:

g097.seq

```

1  ATGGATATTT CAAAACAAAC ATTGCTGGAT AGGGTTTTTA ACCTGAAGGC
51  AAACGGTACG ACGGTACGTA CCGAGTTGAT GGCGGGTTTG ACGACCTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAATC CCCTGATTTT GGGCGAGACC
151 GGAATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CATCCGCCAT
201 CGGCTGTTTT GTCATGGGTT TTATCGGCAA CTATCCGATT GCGCTTGCCC
251 CGGGGATGGG GCTGAATGCC TATTTACCTT TTGCCGTCGT TAAGGGTATG
301 GCGCTGCCTT GGCAGGTGGC GTTGGGTGCG GTGTTTCATT CCGGTCTGAT
351 TTTTCATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCCTATGGG TTTGAAAATG TCGATTGCCG CCGGTATCGG TTTGTTTTTG
451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTCGGC TTGGGCGATA TTCATCAGCC CAGCGCACTG TTGGCATTGT
551 TCGGTTTTGT CATGGTGGTC GTATTGGGGT ATTTCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATTCTGAC GATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAACGAGTTT CACGGCGTGG TCGGCGAAGT ACCGGGCATT GCGCCGACCT
701 TTATGCAGAT GGATTTTAAA GGTCTGTTTA CCGTCAGCAT GGTGAGCGTG
751 ATTTTCGTCT TCTTCTGGT CGATTGTTC GACAGTACCG GAACGCTGGT
801 CGGCGTATCC CACCGTGCCG GACTGCTGGT GGACGGTAAG CTGCCCCGCC
851 TGAACGCGCG ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTTCAACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
951 GGCAGCGGGA CGGACCGGCC TGACGGCGGT TACCGTCGGT GTATTGATGC
1001 TGGCGTGTCT GATGTTCTCC CCATTGGCGA AAAGTGTTC GGTATTTGCC
1051 ACCGCGCCCG CACTGCTTTA TGTCGGCACG CAGATGCTCC GCAGTGCGAG
1101 GGACATTGAT TGGGACGATA TGACTGAAGC CGCGCCCGCG TTCCTGACCA
1151 TTGTCTTCAT GCCGTTTACC TATTCGATTG CAGACGGCAT CGCCTTCGGC
1201 TTCATCAGCT ATGCCGTGGT CAAACTTTTG TGTGCGCCGA CTGGGGACGT
1251 GCCGCTATG GTATGGGTTG TTGCCGTATT GTGGGCATTG AAATTCTGGT
1301 ATTTGGGCTG A

```

This corresponds to the amino acid sequence <SEQ ID 340; ORF 097.ng>:

g097.pep

```

1  MDISKQTLDD RVFNLKANGT TVRTELMAGL TTFLTMCIYV IVNPLILGET
51  GMDMGAVFVA TCIASAI GCF VMGFIGNYP I ALAPMGMLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALPGFVMVV VLGYFRVQGA
201 I I I I L T I T V I A S L M G L N E F H G V V G E V P G I A P T F M Q M D F K G L F T V S M V S V
251 I F V F F L V D L F D S T G T L V G V S H R A G L L V D G K L P R L K R A L L A D S T A I V A G A A
301 L G T S S T T P Y V E S A A G V S A G G R T G L T A V T V G V L M L A C L M F S P L A K S V P V F A
351 T A P A L L Y V G T Q M L R S A R D I D W D D M T E A A P A F L T I V F M P F T Y S I A D G I A F G

```

401 FISYAVVKLL CRRTGDVPPM VWVAVLWAL KFWYLG*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 341>:

```
m097.seq
1  ATGGACACTT CAAAACAAAC ACTGTTGGAC GGGATTTTTA AGCTGAAGGC
51  AAACGGTACK ACGGTGCGTA CCGAGTTGAT GGCGGGTTTG ACAACTTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAACC CTCyGATTTT GGGCGAGACC
151 GGCATGGATA TGGGGGCGGT ATTCGTGCTC ACCTGTATCG CGTCTGCCAT
201 CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
251 CGGGGATGGG GCTGAATGCC TATTTACCTT TGCCGTCGT TAAGGGTATG
301 GGCCTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTTCATCT CCGGTCTGAT
351 TTTTATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTTGTTTTTG
451 GCACTGATT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTTCGGT TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCATTGT
551 TCGGTTTTGC TATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATCTTGAC CATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAATGAATT CACGGCATCA TCGGCGAAGT ACCGAGCATT GCGCCGACTT
701 TTATGCAGAT GGATTTTGAA GGCCTGTTTA CCGTCAGCAT GGTCAAGTGTG
751 ATTTTCGTCT TCTTCTTGGT CGATCTATTT GACAGTACCG GAACGCTGGT
801 CCGCATATCC CACCGTGCCG GGCTGCTGGT GGACGTAAG CTGCCCGCC
851 TGAAACGCGC ACTGCTTGCA GACTCTACCG CCATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTTCACCAC GCCTTATGTG GAAAGCGCGG CGGGCGTATC
951 GGCAGGCGGA CGGACCGGCC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TCGCCTGCCT GATGTTTTCA CCTTTGGCGA AAAGTGTTCC CGCTTTTGCC
1051 ACCGCGCCCG CCTGCTTTA TGTCGGCACG CAGATGCTCC GCAGTGCAG
1101 GGATATTGAT TGGGACGATA TGACGGAAGC CGCACCTGCG TTCCTGACCA
1151 TTGTTTTTCAT GCCGTTTACT TATTCGATTG CAGACGCGAT CGCTTTCGGC
1201 TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCGCA CCAAAGACGT
1251 TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT
1301 ATTTGGGCTG A
```

This corresponds to the amino acid sequence <SEQ ID 342; ORF 097>:

```
m097.pep
1  MDTSKQTLTD GIFKLKANGT TVRTELMAGL TTFLTMCYIV IVNPXILGET
51  GMDMGAVFVA TCIASAIGCF VMGFVGNYP IALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHFRVQGA
201 IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFE GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGIS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VLMLACLMFS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVFMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDVPPM VWIVAVLWAL KFWYLG*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 097 shows 96.3% identity over a 436 aa overlap with a predicted ORF (ORF 097.ng)

from *N. gonorrhoeae*:

```
m097/g097

      10      20      30      40      50      60
m097.pep MDTSKQTLTDGIFKLKANGTTVRTELMAGLTTFLTMCYIVIVNPXILGETGMDMGAVFVA
|| ||||| :||:|||||
g097 MDISKQTLDRVFNKANGTTVRTELMAGLTTFLTMCYIVIVNPLILGETGMDMGAVFVA
      10      20      30      40      50      60

      70      80      90     100     110     120
m097.pep TCIASAIGCFVMGFVGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
|| ||||| :||:|||||
g097 TCIASAIGCFVMGFIGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL
      70      80      90     100     110     120

     130     140     150     160     170     180
```

301

m097.pep	FSFFKVR	EMLVNAL	PMGLKMS	IAAGIGL	FLALISL	KGAGI	IVANPAT	LVGLGDI	HQPSAL
g097	FSFFKVR	EMLVNAL	PMGLKMS	IAAGIGL	FLALISL	KGAGI	IVANPAT	LVGLGDI	HQPSAL
		130	140	150	160	170	180		
m097.pep	LALFGF	FAMVVVL	GHFRVQ	GAIIIT	ILTITV	IASLMGL	NEFHGI	IGEVPS	IAPTFM
g097	LALFGF	FAMVVVL	GHFRVQ	GAIIIT	ILTITV	IASLMGL	NEFHGI	IGEVPS	IAPTFM
		190	200	210	220	230	240		
m097.pep	GLFTV	SMVSVI	FVFFLV	DLFDST	GLVGI	SHRAGL	LLVDGK	LPRLLK	RALLAD
g097	GLFTV	SMVSVI	FVFFLV	DLFDST	GLVGI	SHRAGL	LLVDGK	LPRLLK	RALLAD
		250	260	270	280	290	300		
m097.pep	LGTSS	TPYVES	AAGVS	SAGGRT	GLTAV	TGVLM	LACLMF	SPLAKS	VPFAT
g097	LGTSS	TPYVES	AAGVS	SAGGRT	GLTAV	TGVLM	LACLMF	SPLAKS	VPFAT
		310	320	330	340	350	360		
m097.pep	QMLRS	ARDIDW	DDMTEA	APFLTI	VFMPFT	YSIADG	IAFGFI	SYAVVK	LLCRRT
g097	QMLRS	ARDIDW	DDMTEA	APFLTI	VFMPFT	YSIADG	IAFGFI	SYAVVK	LLCRRT
		370	380	390	400	410	420		
m097.pep	VWIVAV	LWALKF	WYLGX						
g097	VWIVAV	LWALKF	WYLGX						
		430							

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 343>

a097.seq

```

1  ATGGACACTT CAAACAAAC ACTGTGGAC GGGATTTTTA AGCTGAAGGC
51  AAACGGTACG ACGGTGCGTA CCGAGTTGAT GGC GG GTTTG ACAACTTTTT
101 TGACGATGTG CTACATCGTT ATCGTCAACC CTCTGATTTT GGGCGAGACC
151 GGCATGGATA TGGGGGCGGT ATTCGTCGCT ACCTGTATCG CGTCTGCCAT
201 CGGCTGTTTT GTTATGGGTT TTGTCGGCAA CTATCCGATT GCACTCGCAC
251 CGGGGATGGG GCTGAATGCC TATTTCACCT TTGCCGTCGT TAAGGGTATG
301 GGCCTGCCTT GGCAGGTTGC GTTGGGTGCG GTGTTTCATCT CCGGTCTGAT
351 TTTTCATCCTG TTCAGCTTTT TTAAAGTCAG GGAAATGCTG GTCAACGCAC
401 TGCCTATGGG TTTGAAAATG TCGATTGCTG CCGGTATCGG TTTGTTTTTG
451 GCACTGATTT CCCTGAAAGG CGCAGGCATT ATCGTTGCCA ATCCGGCAAC
501 CTTGGTCGGC TTGGGCGATA TTCATCAGCC GTCCGCGTTG TTGGCACTGT
551 TCGGTTTTGC CATGGTGGTC GTATTGGGAC ATTTCCGCGT TCAAGGCGCA
601 ATCATCATCA CCATTTTGAC GATTACCGTC ATTGCCAGCC TGATGGGTTT
651 GAACGAATTT CACGGCATCA TCGCGGAAGT GCCGAGCATT GCGCCGACTT
701 TTATGCAGAT GGATTTTAAA GGGTTGTTTA CCGTCAGCAT GGTGAGCGTG
751 ATTTTCGTCT TTTTCCTAGT CGATCTGTTT GACAGTACCG GAACACTGGT
801 CCGGTGATCG CATCGTGCCG GACTGCTGGT GGACGGTAAG CTGCCCGGCC
851 TGAAACGCGC ACTGCTTGCA GACTCTACCG CTATTGTGGC AGGTGCGGCT
901 TTGGGTACTT CTTCAACCAC GCCTTATGTG GAAAGTGCGG CGGGCGTATC
951 GGCAGGCGGG CGGACAGGTC TGACGGCGGT TACCGTCGGC GTATTGATGC
1001 TCGCCTGCCT GATGTTTTC CTTTGCGCA AAAGTGTTCC CGCTTTTGCC
1051 ACCGCGCCCG CCCTGCTTTA TGTCGGCAGC CAGATGCTCC GCAGTGCGAG
1101 GGACATCGAT TGGGACGATA TGACGGAAGC CGCACCCGCA TTCCTGACCA
1151 TTGCTCTCAT GCCGTTTACC TATTCGATTG CAGACGGCAT CGCTTTCGGC
1201 TTCATCAGTT ATGCCGTGGT TAAACTTTTA TGCCGCCGCA CCAAAGACGT
1251 TCCGCCTATG GTATGGATTG TTGCCGTATT GTGGGCACTG AAATTCTGGT

```

302

1301 ATTTGGGCTG A

This corresponds to the amino acid sequence <SEQ ID 344; ORF 097.a>:

a097.pep

```

1  MDTSKQTLLD GIFKLGANGT TVRTELMAGL TTFLTMCYIV IVNPLILGET
51  GMDMGAVFVA TCIASAIGCF VMGFVGNYP I ALAPGMGLNA YFTFAVVKGM
101 GVPWQVALGA VFISGLIFIL FSFFKVREML VNALPMGLKM SIAAGIGLFL
151 ALISLKGAGI IVANPATLVG LGDIHQPSAL LALFGFAMVV VLGHERVQGA
201 IIITILTITV IASLMGLNEF HGIIGEVPSI APTFMQMDFK GLFTVSMVSV
251 IFVFFLVDLF DSTGTLVGVS HRAGLLVDGK LPRLKRALLA DSTAIVAGAA
301 LGTSSTTPYV ESAAGVSAGG RTGLTAVTVG VMLACLMS PLAKSVPAFA
351 TAPALLYVGT QMLRSARDID WDDMTEAAPA FLTIVMPFT YSIADGIAFG
401 FISYAVVKLL CRRTKDVPPM VWIVAVLWAL KFWYLG*

```

m097/a097 99.3% identity in 436 aa overlap

	10	20	30	40	50	60
m097.pep	MDTSKQTLLDGIFKLGANGTTVRTELMAGLTTFLTMCYIVIVNPIXILGETGMDMGAVFVA					
a097	MDTSKQTLLDGIFKLGANGTTVRTELMAGLTTFLTMCYIVIVNPLILGETGMDMGAVFVA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m097.pep	TCIASAIGCFVMGFVGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL					
a097	TCIASAIGCFVMGFVGNYPIALAPGMGLNAYFTFAVVKGMGVPWQVALGAVFISGLIFIL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m097.pep	FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL					
a097	FSFFKVREMLVNALPMGLKMSIAAGIGLFLALISLKGAGIIVANPATLVGLGDIHQPSAL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m097.pep	LALFGFAMVVVLGHERVQGAIIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFE					
a097	LALFGFAMVVVLGHERVQGAIIITILTITVIASLMGLNEFHGIIGEVPSIAPTFMQMDFK					
	190	200	210	220	230	240
	250	260	270	280	290	300
m097.pep	GLFTVSMVSVIFVFFLVDLFDSTGTLVGISHRAGLLVDGKLPRLKRALLADSTAIVAGAA					
a097	GLFTVSMVSVIFVFFLVDLFDSTGTLVGSHRAGLLVDGKLPRLKRALLADSTAIVAGAA					
	250	260	270	280	290	300
	310	320	330	340	350	360
m097.pep	LGTSSSTTPYVESAGVSAGGRTGLTAVTVGVMLACLMSPLAKSVPAFATAPALLYVGT					
a097	LGTSSSTTPYVESAGVSAGGRTGLTAVTVGVMLACLMSPLAKSVPAFATAPALLYVGT					
	310	320	330	340	350	360
	370	380	390	400	410	420
m097.pep	QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVPPM					
a097	QMLRSARDIDWDDMTEAAPAFLTIVFMPFTYSIADGIAFGFISYAVVKLLCRRTKDVPPM					
	370	380	390	400	410	420
	430					
m097.pep	VWIVAVLWALKFWYLGX					
a097	VWIVAVLWALKFWYLGX					
	430					

```
g098.seq
  1  ATGACCGCCG  ACGGTCTCTT  CGTCGCTTTC  AACTTCAATA  CGTTTGCCGT
51  TGTGCGCAATA  TTGATACCA  TAGCAGCAGG  TGCTGCCCA  CTGGCGCATC
101 AGTTTGTGCG  CGATGTCGCG  CAGTTCGCTG  TCGGGATGGC  TTTGCGCTTC
151 GGGATGAACG  CAGCCGAGCA  TGGACACGCC  GGTACGCATC  ACGTCCATCG
201 GATGGGTATG  TGCAGGCAGG  CTTTCCAAAA  CTTTAATCAC  ACGGATAGGC
251 AGGCCCGCGA  TGGATTGAG  CTTGGTTTTA  TAAGCGGCEA  GCTCGAATTT
301 GTTGGGCAGA  TGGCCGTGAA  TCAGCAAGTG  TGCGACTTCT  TCAAACTCGC
351 ATTTTGTGTC  CAAATTAGAA  TGTGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 346; ORF 098.ng>:

g098.pep

1	MTADGLEVAF	NFNTFAVRI	LIPVQDDAAQ	AGDQFVGDDA	RFAVGMAFAF
51	GMNAAEHGHA	GTHVHRMG	CRQAFQNFH	TDRQAAHGFE	LGFISGQLEF
101	VGQMAVNQQV	CDFFKLAFLC	QIRMS*		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 347>:

```
m098.seq
1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
51  TGTGCGCAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GTTCGGCGATC
101 AGTTTGTFCGG CGATGTCGCG TCGTTCACCT TCCGGATGGC GTTCAAGTTT
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTTCAG CTCTGGTTTTA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTT GGCGACTTCT TCAAACTCGC
351 ATTTTGTGTC CAAATCAGAA TGTGCTAA
```

This corresponds to the amino acid sequence <SEQ ID 348; ORF 098>:

m098.pep

1	MTADGLFVAF	NLNAFAVVRI	LIPVQEDAAE	AGDQFVGDDA	RTFRMAFTF
51	RMNAAQHGYA	GTHYVHRMGM	CRQAFQNFNH	TDRAAHGFE	LGFIISGQLEF
101	VGOMAVNOOV	GDFEKLAFLC	OIRMS*		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 098 shows 89.6% identity over a 125 aa overlap with a predicted ORF (ORF 098.ng) from *N. gonorrhoeae*:

m098/q098

		10	20	30	40	50	60
m098.pep		MTADGLFVAFN	LNAFAVVRIL	IPVQEDAAEAGDQ	FVG	DVARFTFRMAFTFRM	NAAQHGYA
g098		MTADGLFVAFN	FNFTFAVVRIL	IPVQQDAAQAGDQ	FVG	DVARFAVGM	MAFAFGMNAAEHGH
		10	20	30	40	50	60
		70	80	90	100	110	120
m098.pep		GTHYVHRMGMC	RQAFQNFNHTDR	QAAHG	FELGFISGQLEFVG	QMAVNQQVGD	FFKLAFLC
g098		GTHHVHRMGMC	RQAFQNFNHTDR	QAAHG	FELGFISGQLEFVG	QMAVNQQVCD	FFKLAFLC
		70	80	90	100	110	120
m098.pep	QIRMSX						
g098	QIRMSX						

The following partial DNA sequence was identified in *N. meningitidis* <SEO ID 349>:

a098.seq

```
1  ATGACCGCCG ATGGTCTCTT CGTCGCTTTC AACCTCAATG CGTTTGCCGT
51 TGTGCGAATA TTGATACCAG TACAAGAGGA TGCTGCCGAG GCTGGCGATC
```

304

```

101 AGTTTGTCCG CGATGTCGCG CGCTTCACTT TCCGGATGGC TTTCACGTTT
151 AGGATGAACG CAGCCCAGCA TGGATACGCC GGTACGCATT ACGTCCATCG
201 GATGGGTATG TGCAGGCAGG CTTTCCAAAA CTTTAATCAC ACGGATAGGC
251 AGGCCGCGCA TGGATTGAG CTTGGTTTAA TAAGCGGCCA GCTCGAATTT
301 GTTGGGCAGA TGGCCGTGAA TCAGCAGGTG GCGGACTTCT TCAAACCTCG
351 ATTTTGTGTC CAAATCAGAA TGTCGTAA

```

This corresponds to the amino acid sequence <SEQ ID 350; ORF 098.a>:

a098.pep

```

1 MTADGLFVAF NLNAFAVVRI LIPVQEDAAE AGDQFVGDDVA RFTFRMAFTF
51 RMNAAQHGYA GTHYVHRMGM CRQAFQNFNH TDRQAAHGFE LGFISGQLEF
101 VGQMAVNQQV GDFFKLAFLC QIRMS*

```

m098/a098 100.0% identity in 125 aa overlap

	10	20	30	40	50	60
m098.pep	MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVGDDVARFTFRMAFTFRMNAAQHGYA					
a098	MTADGLFVAFNLNAFAVVRI LIPVQEDAAEAGDQFVGDDVARFTFRMAFTFRMNAAQHGYA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m098.pep	GTHYVHRMGMCRQAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQQVGDFFKLAFLC					
a098	GTHYVHRMGMCRQAFQNFNHTDRQAAHGFE LGFISGQLEFVGQMAVNQQVGDFFKLAFLC					
	70	80	90	100	110	120
m098.pep	QIRMSX					
a098	QIRMSX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 351>:

g099.seq

```

1 ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTGGA
51 GCTGACGGGC AAACGGCAGG CGGGCATTAC TGCCACAGAC ATCGTGTGTTG
101 CACTGACCGA ATTCTTGCCT AAAGAGCGCG TGGTCGGGGC GTTTGTGCGAA
151 TTTTTCGGCG AGGGCCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCCATG TTCCGCCATCG
251 ACGCGCAAAAC TATTGATTAT TTGAAACTGA CCGGACGTGA CGACGCGCAG
301 GTGAAATTGG TGGAACCTA CGCCAAAACC GCAGGCTTAT GGGCAGGTGG
351 CTTGAAAACC GCCGTTTATC CGCGCTTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCAGCGCGG TTTTGCCACC
451 GCCGATTGG CGGCGAAAGG GCTGGCGAAG CCTACGAAG AGCCTTCAGA
501 CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCGTGTA
551 CCAATACTTC CAACCCGCGC AACGTTGTCG CCGCCGCACT GTTGGCACGC
601 AATGCCAACC GCCTCGGCTT GAAACGCAAA CCTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGGAATCTA TTTGAAAGAA GCAGGCTTGT
701 TGCCCCGAAAT GGAAAACTC GGCTTCGGTA TCGTCGCCCT CGCATGTACC
751 ACCTGTAACG GCATGAgcgG CGCGCTcgac CCGAAAATCC AACAAGAAAT
801 CATCGACCGC GAtttgtacg cCACCGCCGT ATTGTCAGGC AACCGBAACT
851 TCGACGGCCG TATCCATCCG TATGCGAAAC AGGCTTTTCT CGCTTCGCCCT
901 CCTTTGGTCG TTGCCTACGC ATTGGCAGGT AGCATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGGCCG CGAAATCCGC CTGAAAGATA
1001 TCTGGCCGAC AGACGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051 CCGCAACAAT TCCGCGACAT TTATATCCCG ATGTCCGACA CCGGCACAGC
1101 GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGACCGATG TCCACCTACA
1151 TCCGCCGTCC GCCCTATTGG GAAGGCGCAC TGGCAGGGGA ACGTACATTA
1201 AGAGGTATGC GTCCGCCGGC GATTTTGCCC GACAACATCA CCACCGACCA
1251 CATCTCgcca tCCAATGCGA TTTTGGCCGG cagTGCcgca ggtgaATATT

```

```

1301 TGGCGAAAAT GGGTTTGCTT GAAGAagaCT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTtcgt
1451 tggcacgcgt tgaacCAGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTT ATCATCATTG CCGGTGCGGA
1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC GCGGAAGGGC GTGCGGCTGG
1601 CGGGTGTGGA AGCCATCGCC GCCGAAGGTT TCGAGCGCAT CCACCGCACC
1651 AACCTCATCG GCATGGGCGT CTTGCCGCTG CAATTCAAAC CCGGCACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTGCGCG
1751 AACGCACACC GCGTGCGGC CTGACCCTCG TGATTCACCG TAAAAACGGA
1801 GAAACCGTCG AAGTTCGGT TACCTGCCGC CCCGATACCG CAGAAGAAGC
1851 ATTGGTATAT GAAGCCGGCG GCGTATTGCA ACGGTTTGCA CAGGACTTTT
1901 TGGAAGGGAA CGCGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 352; ORF 099.ng>:

g099.pep

```

1 MLGRASMMRL PDIVGVELTG KRQAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDAQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWAGGLKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAAKGLAK PYEPPSDGQM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLKRRK PWVKSSFAPG SKVAGIYLKE AGLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFILASP
301 PLVVAYALAG SIRFDIENDV LGVADGREIR LKDIWPTDEE IDAIVAEVVK
351 PQQFRDIYIP MSDTGTAQKA PSPLYDWRPM STYIRPPYW EGALAGERTL
401 RGMRPAILP DNITTDHISP SNAILAGSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEPE GQTMRMWEAI
501 ETYMNKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIA AEGFERIHRT
551 NLIGMGVLP QFKPGTNRHT LQLDGTETYD VVGERTPRCG LTLVIHRKNG
601 ETVEVPVTCR PDTAEALVY EAGGVLRFA QDFLEGNA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 353>:

m099.seq

```

1 ATGTGCGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
51 GCTGAACGGC AAACGGCAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
101 CACTGACCGA GTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTGTGCGAA
151 TTCTTCGGCG AGGCGCGGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
251 ATAGCAAAC CATGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
301 GTGAAATTGG TGGAAACCTA CGCCAAACCC GCAGGCTTGT GGGCAGATGC
351 CTTGAAACCC GCCGTTTATC CTCGCGTTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCAAGTAACC CGCATGCCCG TTTTGCGACC
451 GCCGATTGG CCGCGAAAGG GCTGGCGAAG CCTTACGAAG AGCCTTCGGA
501 CGGCCAAATG CCCGACGGCT CGGTCATCAT CGCCGCGATT ACCAGTTGCA
551 CCAACACTTC CAACCCGCGC AACGTTGTTG CCGCCGCGCT CTTGGCAGCG
601 AATGCCAACC GTCTCGGCTT GAAACGCAA CTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTGAAAGAA GCGGGCCTGT
701 TGCCCCGAAAT GGAAGAACTC GGCTTCGGTA TCGTCGCCTT CGCCTGCACC
751 ACCTGCAACG GCATGAGTGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
801 CATCGACCGC GATTGTACG CCACCGCGT ATTATCAGGC AACCGCAACT
851 TCGACGGCCG TATCCACCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901 CCGTTGGTCG TTGCTACGC GCTGGCAGGC AGTATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGCAA GGAAATCCGC CTGAAAGACA
1001 TTTGGCCTGC CGATGAAGAA ATCGATGCCG TCGTTGCCGA ATATGTGAAA
1051 CCGCAGCAGT TCCGCGATGT GTATGTACCG ATGTTTCGACA CCGGCACAGC
1101 GCAAAAAGCA CCCAGTCCGC GTACGATTG GCGTCCGATG TCCACCTACA
1151 TCCGCCGTCC GCCTTACTGG GAAGGCGCGC TGGCAGGGGA ACGCACATTA
1201 AGAGGTATGC GTCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA
1251 CCTCTCGCCG TCCAATGCGA TTTTGGCCGT CAGTGCCGCA GCGAGTATT
1301 TGCGGAAAAT GGGTTTGCTT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGCGGCGACC ACTTGACCGC CCAACGCGCT ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAAAA ACGAAGACGG CAGCGTGC GC CAAGGCTCGT
1451 TCGCCCGCGT CGAACCCGAA GGCGAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGTGCGGA

```

306

```

1551 CTATGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601 CCGGCGTAGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACCTTATCG GCATGGGCGT GTTGCCGCTG CAGTTCAAAC CCGACACCAA
1701 CCGCCATACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTGGTCGGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTCG TGATTCACCG TAAAAACGGC
1801 GAAACCGTTG AAGTTCCCGT TACCTGCTGC CTCGATACTG CAGAAGAAGT
1851 ATTGGTATAT GAAGCCGGCG GCGTGTGCA ACGGTTTGCA CAGGATTTT
1901 TGAAGGGAA CGCGGCTTAG

```

This corresponds to the amino acid sequence <SEQ ID 354; ORF 099>:

```

m099.pep
  1 MLGRASMMRL PDIVGVELNG KRQAGITATD IVLALTEFLR KERVVGAFVE
  51 FFGEGARSLS IGDRTISMN TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
 101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
 151 ADLAAKGLAK PYEEPSDGQM PDGSVIIAAI TSCTNTSNPR NVVAAALLAR
 201 NANRLGLKRK PWVKSSFAPG SKVAEIYLKE AGLLPEMEKL GFGIVAFAC
 251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFILAS
 301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPADEE IDAVVAEYVK
 351 PQQFRDVYVP MFDTGTAQKA PSPLYDWRPM STYIRRPYVW EGALAGERTL
 401 RGMRLAILLP DNITDHLSP SNAILAVSAA GEYLAKMGLP EEDFNSYATH
 451 RGDHLTAQRA TFANPKLFNE MVKNEDGSRV QGSFARVEPE GETMRMWEAI
 501 ETYMNRKQPL IIAAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHR
 551 NLIGMGVLPL QFKPDNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
 601 ETVEVPVTC LDIAEEVLVY EAGGVLRFA QDFLEGNA*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 099 shows 96.2% identity over a 639 aa overlap with a predicted ORF (ORF 099.ng) from *N. gonorrhoeae*:

m099/g099

m099.pep	10	20	30	40	50	60
	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS					
g099	MLGRASMMRLPDIVGVELTGKRQAGITATDIVLALTEFLRKERVVGAFVEFFGEGARSLS					
	10	20	30	40	50	60
m099.pep	70	80	90	100	110	120
	IGDRATISMNTPEFGATAAMFAIDEQTIDYLLKLTGRDDAQVKLVETYAKTAGLWADALKT					
g099	IGDRATISMNTPEFGATAAMFAIDAQTIDYLLKLTGRDDAQVKLVETYAKTAGLWAGGLKT					
	70	80	90	100	110	120
m099.pep	130	140	150	160	170	180
	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGSVIIAAI					
g099	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAKGLAKPYEEPSDGQMPDGAVIDIAAI					
	130	140	150	160	170	180
m099.pep	190	200	210	220	230	240
	TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAEIYLKEAGLLPEMEKL					
g099	TSCTNTSNPRNVVAAALLARNANRLGLKRPWKSSFAPGSKVAGIYLKEAGLLPEMEKL					
	190	200	210	220	230	240
m099.pep	250	260	270	280	290	300
	GFGIVAFACCTTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFILAS					
g099	GFGIVAFACCTTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFILAS					
	250	260	270	280	290	300
	310	320	330	340	350	360

307

m099.pep	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIPADEEIDAVVAEYVKPQQFRDVIYP
g099	PLVVAYALAGSIRFDIENDVLGVADGREIRLKDIPWPTDEEIDAIVA EYVKPQQFRDIYIP
	310 320 330 340 350 360
m099.pep	MFDGTGTAQKAPSPLYDWRPMSYIRRPYWEALAGERTLRGMRPLAILPDNITTDHLS
g099	MSDTGTAQKAPSPLYDWRPMSYIRRPYWEALAGERTLRGMRPPAILPDNITTDHISP
	370 380 390 400 410 420
m099.pep	SNAILAVSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVKNE
g099	SNAILAGSAAGEYLAKMGLPEEDFNSYATHRGDHLTAQRATFANPKLFNEMVRNE
	430 440 450 460 470 480
m099.pep	QGSFARVEPEGETMRMWEAIETYMNKQPLII IAGADYGQSSRDWAAGVRLAGVEAIV
g099	QGSLARVEPEGQTMRMWEAIETYMNKQPLII IAGADYGQSSRDWAAGVRLAGVEAIA
	490 500 510 520 530 540
m099.pep	AEGFERIHRTNLIGMGVLPQFKPDTNRHTLQLDGTETDYDVVGERTPRCDLTLVIHRKNG
g099	AEGFERIHRTNLIGMGVLPQFKPDTNRHTLQLDGTETDYDVVGERTPRCGLTLVIHRKNG
	550 560 570 580 590 600
m099.pep	ETVEVPVTCCLDTAEVLVYEAGGVLRFAQDFLEGNAAX
g099	ETVEVPVTCRPDTAEALVYEAGGVLRFAQDFLEGNAAX
	610 620 630 640

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 355>:

a099.seq

```

1  ATGCTGGGAC GCGCGTCCAT GATGCGCCTG CCCGATATTG TCGGCGTTGA
51  GCTGAACGGC AAACGGAAGG CGGGCATTAC GGCGACGGAT ATTGTGTTGG
101 CACTGACCGA GTTCTGCGC AAAGAACGCG TGGTCGGGGC GTTTGTGCGAA
151 TTCTTCGGCG AGGGCGCGAG AAGCCTGTCT ATCGGCGACC GCGCGACCAT
201 TTCCAACATG ACGCCGGAGT TCGGCGCGAC TGCCGCGATG TTCGCTATTG
251 ATGAGCAAAC CATTGATTAT TTGAAACTGA CCGGACGCGA CGACGCGCAG
301 GTGAAATTGG TGGAACCTA CGCCAAAACC GCAGGCTTGT GGGCAGATGC
351 CTTGAAAACC GCCGTTTATC CGCGCGTTTT GAAATTTGAT TTGAGCAGCG
401 TAACGCGCAA TATGGCAGGC CCGAGCAACC CGCACGCGCG TTTGCGGACC
451 GCCGATTGG CCGGCAAAGG CTTGGCTAAA CCTTACGAAG AGCCTTCAGA
501 CGGCCAAATG CCTGACGGTG CAGTGATTAT TGCCGCGATT ACTTCCTGTA
551 CCAATACTTC CAATCCGCGC AACGTTGTCT CCGCCGCGCT GTTGGCACGC
601 AATGCCAACC GCCTCGGCTT GCAACGCAA CCTTGGGTGA AATCTTCGTT
651 TGCCCCGGGT TCAAAAGTAG CCGAAATCTA TTTGAAAGAA GCAGATCTGC
701 TGCCCCGAAAT GGAAAACTC GGCTTCGGTA TCGTTGCCTT CGCATGTACC
751 ACCTGTAACG GCATGAGCGG CGCGCTGGAT CCGAAAATCC AGAAAGAAAT
801 CATCGACCGC GATTTGTACG CCACCGCCGT ATTGTCAGGC AACCGCAACT
851 TTGACGGCCG TATCCATCCG TATGCGAAAC AGGCTTTCCT CGCTTCGCCT
901 CCGTTGGTCG TTGCCTACGC GCTGGCAGGC AGCATCCGTT TCGATATTGA
951 AAACGACGTA CTCGGCGTTG CAGACGGCAA AGAAATCCGC CTGAAAGACA
1001 TTTGGCCTAC CGATGAAGAA ATCGATGCCA TCGTTGCCGA ATATGTGAAA
1051 CCGCAGCAAT TTCGCGACGT TTATATCCCG ATGTTGCGACA CCGGCACAGC
1101 GCAAAAAGCA CCAAGCCCGC TGTACGACTG GCGTCCAATG TCTACCTATA
1151 TCCGCCGCCC ACCTTACTGG GAAGGCGCAC TGGCAGGGGA ACGCACATTA
1201 AGCGGTATGC GTCCGCTGGC GATTTTGCCC GACAACATCA CCACCGACCA

```

1251 TCTCTCGCCA TCCAATGCGA TTTTGGCAAG CAGTGCCGCA GGCGAATATT
1301 TGGCAAAAAT GGGTTTGCCT GAAGAAGACT TCAACTCTTA CGCAACCCAC
1351 CGTGGCGACC ACTTGACCGC CCAACGCGCA ACCTTCGCCA ATCCGAAACT
1401 GTTTAACGAA ATGGTGAGAA ACGAAGACGG CAGCGTACGC CAAGGTTTCGC
1451 TGGCACGCGT TGAACCCGAA GGCCAAACCA TGCGCATGTG GGAAGCCATC
1501 GAAACCTATA TGAACCGCAA ACAGCCGCTC ATCATCATTG CCGGCGCGGA
1551 CTACGGTCAA GGCTCAAGCC GCGACTGGGC TGCAAAAGGC GTACGCCTCG
1601 CCGGCGTGGA AGCGATTGTT GCCGAAGGCT TCGAGCGTAT CCACCGCACC
1651 AACTTGATCG GTATGGGCGT GTTGCCGCTG CAGTTCAAAC CGGGTACCAA
1701 CCGCCACACC CTGCAACTGG ACGGTACGGA AACCTACGAC GTTGTCGGCG
1751 AACGCACACC GCGCTGCGAC CTGACCCTTG TGATTCACCG TAAAAACGGC
1801 GAGACCGTCG AAGTCCCAT TACCTGCCGC CTCGATACCG CAGAAGAAGT
1851 GTTGGTATAT GAAGCCGTG GCGTATTGCA ACGGTTTGCA CAGGATTTTT
1901 TGGAAGGGAA CCGCGCTTAG

This corresponds to the amino acid sequence <SEQ ID 356; ORF 099.a>:

a099.pep

1 MLGRASMMRL PDIVGVELNG KRKAGITATD IVLALTEFLR KERVVGAFVE
51 FFGEGARSLS IGDRATISNM TPEFGATAAM FAIDEQTIDY LKLTGRDDAQ
101 VKLVETYAKT AGLWADALKT AVYPRVLKFD LSSVTRNMAG PSNPHARFAT
151 ADLAGKGLAK PYEEPSDGM PDGAVIIAAI TSCTNTSNPR NVVAAALLAR
201 NANRLGLQRK PWKSSFAPG SKVAEIYLKE ADLLPEMEKL GFGIVAFAC
251 TCNGMSGALD PKIQKEIIDR DLYATAVLSG NRNFDGRIHP YAKQAFILASP
301 PLVVAYALAG SIRFDIENDV LGVADGKEIR LKDIWPTDEE IDAIVAEYVK
351 PQQFRDVIYP MFDGTGAQKA PSPLYDWRPM STYIRRPYVW EGALAGERTL
401 SGMRLAILP DNITDHLSP SNAILASSAA GEYLAKMGLP EEDFNSYATH
451 RGDHLTAQRA TFANPKLFNE MVRNEDGSVR QGSLARVEFE QQTMRMWEAI
501 ETYMRNRKQPL IIIAGADYGO GSSRDWAAKG VRLAGVEAIV AEGFERIHR
551 NLIGMGVLP QFKPGTNRHT LQLDGTETYD VVGERTPRCD LTLVIHRKNG
601 ETVEVPITCR LDTAEVLVY EAGGVLRFA QDFLEGNA*

m099/a099 97.5% identity in 639 aa overlap

	10	20	30	40	50	60
m099.pep	MLGRASMMRLPDIVGVELNGKRQAGITATDIVLALTEFLRKERVVGAFVEFFFGEGARSLS					
a099	MLGRASMMRLPDIVGVELNGKRKAGITATDIVLALTEFLRKERVVGAFVEFFFGEGARSLS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m099.pep	IGDRATISNMTPEFGATAAMFAIDEQTIDY LKLTGRDDAQVKLVETYAKTAGLWADALKT					
a099	IGDRATISNMTPEFGATAAMFAIDEQTIDY LKLTGRDDAQVKLVETYAKTAGLWADALKT					
	70	80	90	100	110	120
	130	140	150	160	170	180
m099.pep	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAAGLAKPYEEPSDGM PDGSVIIAAI					
a099	AVYPRVLKFDLSSVTRNMAGPSNPHARFATADLAGKGLAKPYEEPSDGM PDGAVIIAAI					
	130	140	150	160	170	180
	190	200	210	220	230	240
m099.pep	TSCTNTSNPRNVVAAALLARNANRLGLKRKPWKSSFAPGSKVAEIYLKEAGLLPEMEKL					
a099	TSCTNTSNPRNVVAAALLARNANRLGLQRKPWKSSFAPGSKVAEIYLKEADLLPEMEKL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m099.pep	GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFILASP					
a099	GFGIVAFACCTCNGMSGALDPKIQKEIIDRDLYATAVLSGNRNFDGRIHPYAKQAFILASP					
	250	260	270	280	290	300
	310	320	330	340	350	360
m099.pep	PLVVAYALAGSIRFDIENDVLGVADGKEIRLKDIWPADEEIDAIVAEYVKPQQFRDVIYP					

g102.seq

This corresponds to the amino acid sequence <SEQ ID 358; ORF 102.ng>:

g102.pap

```

1  MSAKTPSLFG  GAMIIAGKVI  GAGMFPNPTA  NLGDGLIGSL  IVLLYTWFPF
51  SSGALMILEV  NTHNPRGASF  DTMVKDLLGR  GWNIIINGIAV  ALVLYGSTYA
101 YILVGGDLTA  KGIGSAVGGK  ISLTVGQLVF  FGILAFVCWA  SARLVDRFTG
151 VLIGGMVLTF  IWATGGLVAD  AKPSVLFDTQ  APVGTGYWIY  AATALPVCLA
201 SFGFHGNVSS  LLKYFKGDAP  KVAKSIWAGT  LVALVIYVLW  QTAIQSNLPR
251 NEFAPVIAAE  RQLSVLNETL  SKFAQTGDM  KILSLFPYMA  IATSLGVTL
301 GLFDNIADIF  KWNDSMSGGR  TKTVLNFPL  PLISWLLLP  GFFTAIGASG
351 LAATVWDQGI  IPAMLLYVSP  QKIGAGKTYK  VYGGLWMLLV  FLFGIANIAA
401 QVLSQMEVLP  VFKG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 359>:

```

m102.seq
1  ATGCCCAACA  AAACCCCTTC  ACTGTTCCGGC  GCGCGCATGA  TTATCGCCGG
51  CACGGTCATC  GCGCGAGGCA  TGCTCGCCAA  CCCGACCGCC  ACATCCGGCG
101 TATGGTTTAC  CGGCTCGCTG  GCCGTGTTGC  TGTACACCTG  GTTTTCTATG
151 CTTTCCAGCG  GCCTGATGAT  TTTGGAAGTC  AACACCCATT  ATCCGCACGG
201 CGCAAGTTTC  GACACGATGG  TCAAAGACCT  GCTCGGACGC  GGCTGGAACA
251 TCATCAACGG  CATCGCCGTC  GCCTTCGTTT  TATACCTGCT  TACTTACGCT
301 TATATCTTCG  TCGGCGGCGA  CCTGACCGCC  AAAGGCTTAG  GCAGCGCGGC
351 AGGCGGCGAC  GTTCACTCA  CCGTCGGACA  ACTCGTCTTC  TTCGGCATCC
401 TCGCCTTTTG  CGTATGGGCA  TCCGCACGCT  TGGTCGACCG  CTTACCGGC
451 GTCCTTATCG  GCGGCATGGT  ATTGACCTTT  ATTTGGGCGG  CCGGCGGGCT
501 GATTGCCGAT  GCCAAGCCGT  CCGTCTCTTT  CGATACCCAA  GCCCCCGCCG
551 GCACAAACTA  CTGGATTAC  GCCGCCACCG  CCCTGCCCGT  CTGCCTCGCT
601 TCCTTCGGCT  TCCACGGCAA  CGCTCCAGC  CTGCTCAAA  ACTTTAAAG
651 CGACGCGCCC  AAAGTGGCTA  AATCCATCTG  GACGGGCACA  CTGATTGCGC
701 TGGTAATTTA  CGTCCTCTGG  CAAACCGCCA  TCCAAGGCAA  CCTGCCGCGC
751 AACGAGTTTC  CCCCCTCAT  CGCCGCCGAA  GGGCAAGTCT  CCGTCTCAT
801 CGAAACCTTC  TCCAAATTCG  CCCAAACCGG  CAATATGGAC  AAAATATTGT
851 CCCTGTTTTT  CTATATGGCG  ATCGCCACCT  CGTTTTTAGG  CGTAACGCTC
901 GGACTCTTCG  ACTACATCGC  CGACATCTTC  AAATGGAACG  ACAGCATCTC
951 CGGCGCGCAC  AAAACCGCCG  CGCTGACCTT  CCTGCCGCC  CTGATTTCCT
1001 GCCTGCTCTT  CCCACCGGCG  TTCGTTACCG  CCATCGGCTA  CGTCGGCCTG
1051 GCGGCAACCG  TCTGGACAGG  CATCATCCCC  GCCATGCTGC  TCTACCGTTC
1101 GCGCAAAAAA  TTCGGCGCAG  GCAAAACCTA  TAAAGTTTAC  GCGCGCTTGT
1151 GGCTGATGGT  TTGGGTCTTC  CTTTTCGGCA  TCGTCAACAT  CGCCGCACAG
1201 GTATTGAGCC  AAATGGAAC  CGTCCCCGTA  TTTAAAGGAT  AA

```

This corresponds to the amino acid sequence <SEQ ID 360; ORF 102>:

```

m102.pep..
1  MPNKTPSLFG  GAMIIAGTVI  GAGMLANPTA  TSGVWFTGSL  AVLLYTWFSM
51  LSSGLMILEV  NTHYPHGASF  DTMVKDLLGR  GWNIIINGIAV  AFVLYLLTYA
101 YIFVGGDLTA  KGLGSAAGGD  VSLTVGQLVF  FGILAFVCWA  SARLVDRFTG
151 VLIGGMVLTF  IWAAGGLIAD  AKPSVLFDTQ  APAGTNYWIY  AATALPVCLA
201 SFGFHGNVSS  LLKYFKGDAP  KVAKSIWGT  LIALVIYVLW  QTAIQGNLPR
251 NEFAPVIAAE  GQSVLIETL  SKFAQTGNMD  KILSLFSYMA  IATSLGVTL
301 GLFDYIADIF  KWNDSISGRT  KTAALTFLPP  LISCLLPPTG  FVTAIGYVGL
351 AATVWTGIIP  AMLLYRSRKK  FGAGKTYKVY  GGLWLMVWVF  LFGIVNIAAQ
401 VLSQMEVLPV  FKG*

```

m102/g102 86.0% identity in 415 aa overlap

	10	20	30	40	50	60
m102.pep	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
g102	MSAKTPSLFGGAMIIAGKVIAGMFPNPTANLGDGLIGSLIVLLYTWFPFSSGALMILEV					
	10	20	30	40	50	60
m102.pep	NTHYPHGASFDTMVKDLLGRGWNIIINGIAVAFVLYLLTYAIFVGGDLTAKGLGSAAGGD					
g102	NTHNPRGASFDTMVKDLLGRGWNIIINGIAVALVLYGSTYAYILVGGDLTAKGIGSAVGGK					
	70	80	90	100	110	120
m102.pep	VSLTVGQLVFFGILAFVCWASARLVDRFTGVLLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
g102	ISLTVGQLVFFGILAFVCWASARLVDRFTGVLLIGGMVLTFIWTGGLVADAKPSVLFDTQ					
	130	140	150	160	170	180
m102.pep	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGLIALVIYVLW					
g102	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGLIALVIYVLW					

a102.seq

1	ATGCCACCA	AAACCCCTTC	ACTGTTCGGC	GGCGCATGA	TTATCGCCGG
51	CACGNTCATC	GGCGCAGSTA	TGCTGCACAA	CCCACCGCC	ACATCCGGCG
101	TATGGTTTAC	CGGCTCGCTG	GCCGTGTTGC	TGTACCACTG	GTTTTCATGT
151	CTCTCCAGCG	GCCTGATGAT	TTTGGAACTA	AACACCCACT	ACCCCCACGG
201	CGCGANCTTC	GACACCATGG	TTAAAGACCT	GCTCGGACGG	AGCTGGAACA
251	TCATCAACGG	CATCGCCGTC	GCCCTCGTTT	TATACCTGCT	TACTTACGCT
301	TATATCTTCG	CTCGCGCGCA	CGTCAGCCGC	AAAGGCTTAG	CGACGCGCGC
351	AGGCGGCAAT	GTTTCACTCA	CCGTCCGACA	ACTCGTCTTC	TTCGGCATTC
401	TCGCCTTTTG	CGTATGGGCA	TCCGCACGCT	TGGTCGACCG	ATTCAACAGC
451	GTCCTCATCG	CGGGCATGGT	ATTAACCTTT	ATTTGGGCAA	CCGGCGGCCT
501	GATTGCCGAT	GCCAAACTGC	CGTCCTCTTT	CGACACCCAA	CGCCCTACCG
551	GCACCAACTA	CTGGATTAT	GTCGCCACCG	CCCTGCCCGT	CTGCCTTGCG
601	TCATTCGGTT	TCCACGGCAA	CGTCTCCAGC	CTGCTCAAT	ACTTTAAAGG
651	CGACGCGCCC	AAAGTGCTTA	AATCAACTCTG	CAGCGGCACA	CTGATTGCGC
701	TGGTAATTTA	CGTCTCTGG	CAACACGCCA	TCCAANGCAA	CTCTGCCGCGC
751	AACGAGTTTC	CCCCCGTGAT	TGCCGCCGAA	GGGCAAGTCT	CCGTCTNTGAT
801	TGAAACCCCT	TCCAAATTCG	CCCCAACCGG	CAATATGGAC	AAAATATTGT
851	CCCTGTTTTC	CTATATGGCG	ATCGCCACCT	CGTTTITAGG	CGTAACGCTG
901	GGACTCTTCG	ACTACATCGC	GCACATCTTC	AAATGGAACG	ACAGCGTGTC
951	CGGCGCGCCT	AAAACCCGCG	CGCTGACCTT	CCTGCCGCCCT	NTAATTTTCT
1001	GCCTGCTCTT	CCCCACCGCG	TTTGTTACCG	CCATCGGNTA	CGTCGGCGCTG
1051	GCGGCAACCG	TCTGACACAG	CATCATCCCC	GCATGCTGTC	TNTACCGTTC
1101	GCGCAAAAAA	TTCGGCGCAG	CAAAAACCTA	TAAAGTTTAC	GGCGGCTTGT
1151	GGCTGATGGT	TTGGGTCTTC	CTTTTCGGCA	TCNTCAACAT	CGCCGCACAN
1201	GTATTGAGCC	AAATGGAAC	CGTCCCCGTA	TTTAAAGGAT	AA
1202					

a102.pep

1	MPTKTPSLFG	GAMIAGTXI	GAGMLANPTA	TSGVWFTGSL	AVLLYTWFMS
51	LSSGGMILEV	NTHYPHGAXF	DTMVKDLGR	SWNIINGIAV	AFVLYLLTYA
101	YIFVGGDLTA	KGLSAAGGN	VSLTVGQLVF	FGILAFVCWA	SARLVDRFST
151	VLIGGGMVLT	IWATGGLDA	AKLPVLFDTQ	APTGTNYWYI	VATALPVLCA
201	SFGFHGNVSS	LLKYFKGDAP	KVAKSIWTGT	LIALVIYVLW	QTAIPXNLPR
251	NEFAPVIAAE	GQVSXVIETL	SKFAQTGNMD	KILSLFSYMA	IATSFLGVTL
301	GLFDYIADIF	KWSDSVSGRT	KTAALTFLLP	XISCLLFPTG	FVTAIGYVGL
351	AATVWGTGI	PMLLYRSRKK	FGAGKTYKVY	GGLWLMVWVF	LFGIXNIAAX
401	VLSQMLVPV	FKG*			

m102 / a102 95.9% identity in 413 aa overlap

312

m102.pep	10	20	30	40	50	60
	MPNKTPSLFGGAMIIAGTVIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
a102	MPTKTPSLFGGAMIIAGTXIGAGMLANPTATSGVWFTGSLAVLLYTWFSMLSSGLMILEV					
	10	20	30	40	50	60
m102.pep	70	80	90	100	110	120
	NTHYPHGASFDTMVKDLLGRGWNIIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGD					
a102	NTHYPHGAXFDTMVKDLLGRSWNIIINGIAVAFVLYLLTYAYIFVGGDLTAKGLGSAAGGN					
	70	80	90	100	110	120
m102.pep	130	140	150	160	170	180
	VSLTVGQLVFFGILAFVCWASARLVDRFTGVLIGGMVLTFIWAAGGLIADAKPSVLFDTQ					
a102	VSLTVGQLVFFGILAFVCWASARLVDRFTSVLIGGMVLTFIWAAGGLIADAKLPVLFDTQ					
	130	140	150	160	170	180
m102.pep	190	200	210	220	230	240
	APAGTNYWIYAATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW					
a102	APTGTNYWIYVATALPVCLASFGFHGNVSSLLKYFKGDAPKVAKSIWTGTLIALVIYVLW					
	190	200	210	220	230	240
m102.pep	250	260	270	280	290	300
	QTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
a102	QTAIQXNLPRNEFAPVIAAEGQVSVXIETLSKFAQTGNMDKILSLFSYMAIATSFLGVTL					
	250	260	270	280	290	300
m102.pep	310	320	330	340	350	360
	GLFDYIADIFKWNDSISGRKTAALTFLPPLISCLLEPTGFVTAIGYVGLAATVWTGIIP					
a102	GLFDYIADIFKWNDSVSGRKTAAALTFLPPXISCLLEPTGFVTAIGYVGLAATVWTGIIP					
	310	320	330	340	350	360
m102.pep	370	380	390	400	410	
	AMLLYRSRKKFAGKTYKVYGGGLWLMVWVFLFGIVNIAAQVLSQMELVPVFKGX					
a102	AMLLYRSRKKFAGKTYKVYGGGLWLMVWVFLFGIXNIAAXVLSQMELVPVFKGX					
	370	380	390	400	410	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 363>:

g105.seq

1	Atgtccgcag	aaaCATACac	acAAAtcggc	tGGgtaggct	taggGcaa
51	gGgtctgcct	atgGTAACGC	GGCTCTTGA	CGGCGGCATC	GAAGTCGGCG
101	TATACAACCG	CTCGCCCGAC	AAAAC TGCCC	CCATCTCgcg	CAAAGGAGCA
151	AAAGTTTACG	GCagcACCGC	CGAACTCGTC	CGCGCTGCC	CCGTCATTTT
201	CCTGATGGTT	TCCGACTATG	CCGCCGTGTG	CGACATCCTG	AACGGAGTCC
251	GCGACGGATT	GGCCGGCAAA	ATCATCGTCA	ACATGAGCAC	CATCTCCCCG
301	ACCGAAAACC	TCGCCGTCAA	AGCACTTGTC	GAAGCCGCAG	GCGGACAGTT
351	TGCCGAAGCA	CCCGTTTCCG	GATCGGTCGG	ACCGCCACC	AACGGCACAC
401	TGCTGATTCT	GTTCGGCGGC	AGCGAAGCCG	TTTTAAACCC	GCTGCAAAAA
451	ATATTTTCCC	TTGTCGGCAA	AAAAACCTTC	CATTTCTGGCG	ATGTCGGCAA
501	AGGCTCGGGC	GCGAAACTCG	TCTTGAACCT	GCTCTTAGGC	ATTTTCTGGCG
551	AAGCGTACAG	CGAAGCGATG	CTGATGGCGC	GGCAGTTCCG	CATCGATACC
601	GACACCATCG	TCGAAGCCAT	CGGCGGCTCG	GCAATGGACT	CGCCTATGTT
651	TCAAACAAAA	AAATCACTAT	GGGCAAACCG	TGAGTTCCCC	CCTGCCTTTG
701	CACTCAAACA	CGCTTCCAAA	GAcctTAACC	TCGccgtcAA	AGAGCTTGAA
751	CAGGCAGGCA	ACACCCTGCC	CGCCGTCGAA	ACCGTTGCTG	CCAGCTACCG
801	CAAAGCAGTT	GAAAGCCGGCT	ACGGCGAACA	GGACGTTTCC	GGCGTTTACC
851	TGAAATTGGC	AGAACACTGA			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 365>:

This corresponds to the amino acid sequence <SEQ ID 366; ORF 105>:

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 105 shows 79.9% identity over a 289 aa overlap with a predicted ORF (ORF 105.ng) from *N. gonorrhoeae*:

m105/g105

		10	20	30	40	50	60
g105.pep		MSAETYYTQIGWVVLGQMGLPMVTRLRLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV					
		: : : :					
m105		MSANEYAQIGWIGLQGMGLPMVTRLRLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
		10	20	30	40	50	60
		70	80	90	100	110	120
g105.pep		RACPVIIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA					
		:					
m105		RDYPVIIFLMVSDYAAVCDILNGVRDGLAGXIIIVNMSTISPTKLAVKALVEAQR-QFAEA					
		70	80	90	100	110	
		130	140	150	160	170	180
g105.pep		PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKSGAKLVLSNLLG					
		: : : : : : : :					
m105		PVSGSVGPATNGTLLILFGGSEPFXTCKKYFPSSAKKP-SISAMSAKVRARNSSXTRSW					
		120	130	140	150	160	170

```

              190      200      210      220      230      240
g105.pep    IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
             |:  :: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
m105        AFSANVQRXXLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPXXFALKHASK
            180      190      200      210      220      230

              250      260      270      280      289
g105.pep    DLNLAVKELEQAGNTLPAVETVAASyrKAVEAGYGEQDVSGVYLKLAEH
             ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
m105        DLNLAVKELEQAGNTLPAVETVAASyrKAVEAGYGTQDVSGVYLKLAEH
            240      250      260      270      280

```

a105.seq

1	ATGTCGCCAA	ACGAATACAC	ACAAATCGGC	TGGATAGGCT	TAGGGCAAT
51	GGGTCTGCCT	ATGGTAACGC	GGCTCTTGGA	CGGCGGCATC	GAAGTCGGCG
101	TATACAAACCG	CTCGCCCGAC	AAAATCGGCC	CCATCTCCGC	CAAAGCGCCA
151	AAAGTTTACG	GCAACACCGC	CGAACTCGTC	CGCGACTATC	CCGTCATTTT
201	CTCGATGGTT	TCCGACTATG	CGCCCGTGTG	CGACATCTCG	AACGGAGTCC
251	GCGACGGATT	GGCCGGCAAA	ATCATCGTCA	ACATGAGCAC	CATCTCCCCG
301	ACCGAAAAAC	TCGCCGTCAA	AGCATCTTGC	GAAGCCGCAG	GCGGCAGGTT
351	TGCCGAAGCA	CCCGTTTCCG	GATCGGTCGG	GCCCGCCACC	AACGGCACGC
401	TGCTGATTCT	TCTCGCGCGC	AGCGAAGCCG	TTTTAAACCC	GCTGCAAAAA
451	ATATTTTCCC	TCGTGCGCAA	AAAAACCTTC	CATTTGCGCG	ATGTCGCGCA
501	AGGTTGCGGC	GCGAAACTCG	TCTTGAACTC	GCTCTTGGGC	ATTTTCGCGC
551	AAGCGTACAG	CGAAGCGATG	CTGATGGCTG	GCGAGTTCGG	CATCGATACC
601	GACACCATCG	TCGAAGCCAT	CGGCGGCGCG	CGAATGGACT	CGCCATGTTT
651	CAAAACCAAA	AAATCCTGTG	GGGCAAAACG	CGAATCCCA	GCCGCGTTCC
701	CCCTCAAACA	CGCCTCCAAA	GACCTCAACC	TCGCCGTCAA	AGAGCTTGAA
751	CAGGACGGCA	ACACCTTGCC	CGCCGTCGAA	ACGTTGTGTG	CCAGCTACCG
801	CAAGACGATC	GAAGCCGGCT	ACGGCGAACA	GGACGTTTCC	GGCGTTTACC
851	TGAAATTGGC	AGAACTACTG			

a105.pep

1	MSANEYTRQIG	WIGLGQMGLP	MVTRLLDGGI	EVGVYNRSPD	KTAPISAKGA
51	KVYGNATLAL	RDYPVIFLMV	<u>SDYAAVCDIL</u>	NGVRDLGLAK	IIVNMTSTSP
101	TENLVAELV	EAAAGQFAEA	PSVSGVGPAT	NGTLLILFGG	SEAMNLPLQK
151	IFSLVGKKTf	HFGDVGKGSg	AKLVlNSLLg	IFGEAYSEAM	LMARQFGIDT
201	DTIVEAIGGS	AMDSPMFQTK	KSLWANREFP	PAPALKHAASK	DLNLAVKELE
251	OAGNTLPAVE	TVAASYRKAV	EAGYGEODFS	GUYLKLAHS*	

m105/a105 96.5% identity in 289 aa overlap

		10	20	30	40	50	60	
m105.pep		MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV						
		:						
a105		MSANEYTQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV						
		10	20	30	40	50	60	
		70	80	90	100	110	119	
m105.pep		RDYPVIFLMSVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPENLAVKALVEAAG-QFAEA						
a105		RDYPVIFLMSVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPENLAVKALVEAAGGQFAEA						
		70	80	90	100	110	120	
	120	130	140	150	160	170	179	
m105.pep		PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSGAKLVLSLLG						
a105		PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGSGAKLVLSLLG						
		130	140	150	160	170	180	
	180	190	200	210	220	230		

315

```

m105.pep  IFGDV-QRXMLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFFXAFALKHASK
           |||:: :| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a105       IFGEAYSEAMLMARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFFPAFALKHASK
           190      200      210      220      230      240

           240      250      260      270      280
m105.pep  DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
           ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
a105       DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAEHX
           250      260      270      280      290

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 369>:

```

g105-1.seq
1  ATGTCGCGCAG  AAACATACAC  ACAAATCGGC  TGGGTAGGCT  TAGGGCAAAT
51  GGGTCTGCCT  ATGCTAACGC  GGCTCTTGA  CGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCCGAC  AAAACTGCC  CCATCTCCGC  CAAAGGAGCA
151 AAAGTTTACG  GCAGCACC GC  CGAACTCGTC  CGCGCCTGCC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CGGCCGTGTG  CGACATCTGT  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCGG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCCGTTTCCG  GATCGGTGCG  ACCCGCCACC  AACGGCACAC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TGTTCGGCAA  AAAACCTTC  CATTTCGGCG  ATGTCGGCAA
501 AGGCTCGGGC  GCGAAACTCG  TCTTGAACTC  GCTCTTAGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAGCGATG  CTGATGGCGC  GGCAGTTCGG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGCGGCTCG  GCAATGGACT  CGCCTATGTT
651 TCAAAACAAA  AAATCACTAT  GGGCAAACCG  TGAGTTCCCG  CCTGCCTTTG
701 CACTCAACA  CGCTTCCAAA  GACCTTAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCTTGGC  CGCCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTC  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GCGGTTTACC
851 TGAATTGGC  AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 370; ORF 105-1.ng>:

```

g105-1.pep
1  MSAETTYQIG  WVGLGQMLP  MVTRLDDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGYSTAEIV  RACPVIFLMV SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP
101 TENLAVKALV  EAAGGQFAEA  PVSGSVGPAT NGTLILFGG  SEAVLNPLQK
151 IFSLVGKKTF  HFGDVGKSG  AKLVLSLLG  IFGEAYSEAM  LMARQFGIDT
201 DTIVEAIGGS  AMDSPMFQTK  KSLWANREFP  PAFALKHASK  DLNLAVKELE
251 QAGNTLPAVE  TVAASYRKAV  EAGYGEQDVS  GVYLKLAEH*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 371>:

```

m105-1.seq
1  ATGTCGCGCAA  ACGAATACGC  ACAAATCGGC  TGGATAGGCT  TAGGGCAAAT
51  GGGTCTGCCT  ATGCTAACGC  GGCTCTTGA  CGCGGCATC  GAAGTCGGCG
101 TATACAACCG  CTCGCCCGAC  AAAACTGCC  CCATCTCCGC  CAAAGGCGCA
151 AAAGTTTACG  GCAACACCGC  CGAACTCGTC  CGCGACTATC  CCGTCATTTT
201 CCTGATGGTT  TCCGACTATG  CGGCCGTGTG  CGACATCTGT  AACGGAGTCC
251 GCGACGGATT  GGCCGGCAA  ATCATCGTCA  ACATGAGCAC  CATCTCCCGG
301 ACCGAAAACC  TCGCCGTCAA  AGCACTTGTC  GAAGCCGCAG  GCGGACAGTT
351 TGCCGAAGCA  CCCGTTTCCG  GATCGGTGCG  GCCCGCCACC  AACGGCACGC
401 TGCTGATTCT  GTTCGGCGGC  AGCGAAGCCG  TTTTAAACCC  GCTGCAAAAA
451 ATATTTTCCC  TCGTCGGCAA  AAAACCTTC  CATTTCGGCG  ATGTCGGCAA
501 AGGTTTCGGC  GCGAAACTCG  TCTTGAACTC  GCTCTGGGC  ATTTTCGGCG
551 AAGCGTACAG  CGAAnCGATG  CTGATGGCGC  GGCAGTTCGG  CATCGATACC
601 GACACCATCG  TCGAAGCCAT  CGGsGACTCG  GCAATGGACT  CGCCCATGTT
651 CCAAACCAA  AAATCCCTGT  GGGCAAACCG  CGAATTCCCG  CCCGCCTTCG
701 CCCTCAACA  CGCCTCCAAA  GACCTCAACC  TCGCCGTCAA  AGAGCTTGAA
751 CAGGCAGGCA  ACACCTTGGC  CGCCGTCGAA  ACCGTTGCTG  CCAGCTACCG
801 CAAAGCAGTC  GAAGCCGGCT  ACGGCGAACA  GGACGTTTCC  GCGGTTTACC
851 TGAACCTGGC  AGAACACTGA

```

This corresponds to the amino acid sequence <SEQ ID 372; ORF 105-1>:

```

m105-1.pep
1  MSANEYAQIG  WIGLGQMLP  MVTRLDDGGI  EVGVYNRSPD  KTAPISAKGA
51  KVGYNTAEIV RDYPVIFLMV SDYAAVCDIL  NGVRDGLAGK  IIVNMSTISP

```

316

101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
 151 IFSLVGKKTFF HFGDVGKGS AKLVNLSLLG IFGEAYSEX LMARQFGIDT
 201 DTIVEAIGDS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
 251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*

m105-1/g105-1 96.9% identity in 289 aa overlap

	10	20	30	40	50	60
m105-1.pep	MSANEYAQIGWIGLQGMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
g105-1	MSAETYTQIGWVGLGQGMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGSTAELV					
	10	20	30	40	50	60
m105-1.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPENLAVKALVEAAGGQFAEA					
g105-1	RACPVIFLMVSDYAAVCDILNGVRDGLAGKIIVNMSTISPENLAVKALVEAAGGQFAEA					
	70	80	90	100	110	120
m105-1.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGS AKLVNLSLLG					
g105-1	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVGKGS AKLVNLSLLG					
	130	140	150	160	170	180
m105-1.pep	IFGEAYSEXLMARQFGIDTDTIVEAIGDSAMDSPMFQTKKSLWANREFPPAFALKHASK					
g105-1	IFGEAYSEAMLARQFGIDTDTIVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK					
	190	200	210	220	230	240
m105-1.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVS GVYLKLAEHX					
g105-1	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVS GVYLKLAEHX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 373>:

a105-1.seq
 1 ATGTCCGCAA ACGAATACAC ACAAATCGGC TGGATAGGCT TAGGGCAAAT
 51 GGGTCTGCCT ATGGTAACGC GGCTCTTGA CCGCGGCATC GAAGTCGGCG
 101 TATACAACCG CTCGCCGAC AAAACTGCCC CCATCTCCGC CAAAGGCGCA
 151 AAAGTTTACG GCAACACCGC CGAACTCGTC CGCGACTATC CCGTCATTTT
 201 CCTGATGGTT TCCGACTATG CCGCCGTGTG CGACATCCTG AACGGAGTCC
 251 GCGACGGATT GGCCGGCAAA ATCATCGTCA ACATGAGCAC CATCTCCCGC
 301 ACCGAAACC TCGCCGTCAA AGCACTTGTC GAAGCCGCAG GCGGACAGTT
 351 TGCCGAAGCA CCCGTTTCCG GATCGGTCCG GCCCGCCACC AACGGCACGC
 401 TGCTGATTCT GTTCGGCGGC AGCGAAGCCG TTTTAAACCC GCTGCAAAAA
 451 ATATTTTCCC TCGTCGGCAA AAAAACCTTC CATTTCGGCG ATGTCGGCAA
 501 AGGTTCCGGC GCGAAACTCG TCTTGAACCT GCTCTTGGGC ATTTTCGGCG
 551 AAGCGTACAG CGAAGCGATG CTGATGGCGC GGCAGTTCGG CATCGATACC
 601 GACACCATCG TCGAAGCCAT CGGCGGCTCG GCAATGGACT CGCCCATGTT
 651 CCAACCAAAA AATCCCTGT GGGCAAACCG CGAATTCCCA CCCGCTTCG
 701 CCCTCAAACA CGCCTCCAAA GACCTCAACC TCGCCGTCAA AGAGCTTGAA
 751 CAGGCAGGCA ACACCTGCC CGCCGTGCAA ACCGTTGCTG CCAGCTACCG
 801 CAAAGCAGTC GAAGCCGGCT ACGGCGAACA GGACGTTTC GCGCTTACG
 851 TGAATTGGC AGAACACTGA

This corresponds to the amino acid sequence <SEQ ID 374; ORF 105-1.a>:

a105-1.pep
 1 MSANEYTQIG WIGLQGMGLP MVTRLLDGGI EVGVYNRSPD KTAPISAKGA
 51 KVGNTAELV RDYPVIFLMV SDYAAVCDIL NGVRDGLAGK IIVNMSTISP
 101 TENLAVKALV EAAGGQFAEA PVSGSVGPAT NGTLLILFGG SEAVLNPLQK
 151 IFSLVGKKTFF HFGDVGKGS AKLVNLSLLG IFGEAYSEAM LMARQFGIDT
 201 DTIVEAIGDS AMDSPMFQTK KSLWANREFP PAFALKHASK DLNLAVKELE
 251 QAGNTLPAVE TVAASYRKAV EAGYGEQDVS GVYLKLAEH*

a105-1/m105-1 99.0% identity in 289 aa overlap

	10	20	30	40	50	60
a105-1.pep	MSANEYTQIGWIGLQGMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNTAELV					
	: : : : : :					

317

m105-1	MSANEYAQIGWIGLGQMGLPMVTRLLDGGIEVGVYNRSPDKTAPISAKGAKVYGNATAEIV
	10 20 30 40 50 60
a105-1.pep	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
m105-1	RDYPVIFLMVSDYAAVCDILNGVRDGLAGKIIIVNMSTISPTENLAVKALVEAAGGQFAEA
	70 80 90 100 110 120
a105-1.pep	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVKGSGAKLVLSLLG
m105-1	PVSGSVGPATNGTLLILFGGSEAVLNPLQKIFSLVGKKTFFHFGDVKGSGAKLVLSLLG
	130 140 150 160 170 180
a105-1.pep	IFGEAYSEAMLMARQFGIDTDITVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
m105-1	IFGEAYSEAMLMARQFGIDTDITVEAIGGSAMDSPMFQTKKSLWANREFPPAFALKHASK
	190 200 210 220 230 240
a105-1.pep	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAHX
m105-1	DLNLAVKELEQAGNTLPAVETVAASYRKAVEAGYGEQDVSGVYLKLAHX
	250 260 270 280 290

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 375>:

g107.seq

```

1  ATGGTATTAA CCTTTATTG GGCACCGGC GGCCTGGTTG CCGATGCCAA
51  ACGGTCGGTC CTCTTCGACA CCCAAGCCCC CGTCGGCACC GGCTACTGGA
101 TTTACGCCGC CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTTT CCAGCCTGCT CAAATACTTT AAAGGCGACG cgcCcaaagt
201 GgCGAaATCc atctggGcag gtacattggt TGCCttggtta atttacgtcc
251 TCTggcaaac cgccatCcaa agcaacctGC cgcgcaacga gttcgcCCCCc
301 gtgattgccg ccgagaggca actCTCCGTC CTgaatgaaa cccTgtccaa
351 attcgcccaa accggcgata tggataAaat attgtcccta tttccctaca
401 tggcaatcgc cacctccttt ttaggcgTAA Ccttaggcct gtttgacaac
451 atcgccggac atcttcaaat ggaacgacag tatgtccggg cggcaccaaa
501 accgtcgcgc tga

```

This corresponds to the amino acid sequence <SEQ ID 376; ORF 107.ng>:

g107.pep

```

1  MVLTFIWATG GLVADAKPSV LFDTPQAPVGT GYWIYAATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWAGTLVALV IYVLWQTAIQ SNLPRNEFAP
101 VIAAERQLSV LNETLSKFAQ TGDMDKILSL FPYMAIATSF LGVTLGLFDN
151 IAGHLQMERQ YVRAAPKPSR *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 377>:

m107.seq

```

1  ATGGTATTGA CCTTTATTG GCGGCCCGGC GGCCTGATTG CCGATGCCAA
51  GCCGTCCGTC CTCTTCGATA CCCAAGCCCC CGCCGGCACA AACTACTGGA
101 TTTACGCCGc CACCGCCCTG CCCGTCTGCC TCGCTTCCTT CGGCTTCCAC
151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TCGCTGGTA ATTTACGTCC
251 TCTGGCAAAc CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCCc
301 GTCATCGCCG CCGAAGGGCA AGTCTCCGTC CTCATCGAAA CCCTGTCCAA
351 ATTGCCCCAA ACCGGCAATA TGGACAAAT ATTGTCCCTG TTTTCCTATA
401 TGGGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCCATC TTCAAATGGA ACGACAGCAT CTCCGGgCCG CACCAAAACC
501 GCCGCGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 378; ORF 107>:

m107.pep..

```

1  MVLTFIWAAG GLIADAKPSV LFDTPQAPAGT NYWIYAXTAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP

```

101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY
151 IAHLMERQH LRAAPKPPR*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 107 shows 89.4% identity over a 170 aa overlap with a predicted ORF (ORF 107.ng)
from *N. gonorrhoeae*:

m107/g107

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	: : : : :					
g107	MVLTFIWATGGLVADAKPSVLFDTQAPVGTGYWIYAATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60
	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
	: : : : :					
g107	KGDAPKVAKSIWAGTLVALVIYVLWQTAIQSNLPRNEFAPVIAAERQLSVLNETLSKFAQ					
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSF LGVTLGLFDYIA-HLMERQHLRAAPKPPR					
	: : : : :					
g107	TGDMDKILSLFPYMAIATSF LGVTLGLFDNIAGHLQMERQYVRAAPKPSR					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 379>:

a107.seq

```

1  ATGGTATTAA CCTTTATTG GGCACCGGC GGCCTGATTG CCGATGCCAA
51  ACTGCCCGTC CTCTTCGACA CCCAAGCCCC TACCGGCACC AACTACTGGA
101 TTTATGTCGC CACCGCCCTG CCCGTCTGCC TTGCGTCATT CGGTTTCCAC
151 GGCAACGTCT CCAGCCTGCT CAAATACTTT AAAGGCGACG CGCCCAAAGT
201 GGCTAAATCC ATCTGGACGG GCACACTGAT TGGCTGGTA ATTACGTCC
251 TCTGGCAAAC CGCCATCCAA GGCAACCTGC CGCGCAACGA GTTCGCCCCC
301 GTGATTGCCG CCGAAGGGCA AGTCTCCGTC CTGATTGAAA CCCTGTCCAA
351 ATTCGCCCAA ACCGGCAATA TGGACAAAT ATTGTCCCTG TTTTCCTATA
401 TGGCGATCGC CACCTCGTTT TTAGGCGTAA CGCTCGGACT CTTCGACTAC
451 ATCGCCGACA TCTTCAAATG GAACGACAGC GTGTCCGGCC GCACCAAAAC
501 CGCCGCGCTG ACCTTCCTGC CGCCTCTAAT TTCCTGCCTG CTCTTCCCA
551 CCGGCTTTGT TACCGCCATC GGCTACGTCG GCCTGGCGGC AACCGTCTGG
601 ACAGGCATCA TCCCGCCCAT GCTGCTCTAC CGTTCGCGCA AAAAATTCGG
651 CGCAGGCAA ACCTATAAAG TTTACGGCGG CTTGTGGCTG ATGGTTTGGG
701 TCTTCCTTTT CGGCATCGTC AACATCGCCG CACAGGTATT GAGCCAAATG
751 GAACTCGTCC CCGTATTATA AGGATAA

```

This corresponds to the amino acid sequence <SEQ ID 380; ORF 107.a>:

a107.pep

```

1  MVLTFIWATG GLIADAKLPV LFDQTAPTGT NYWIYVATAL PVCLASFGFH
51  GNVSSLLKYF KGDAPKVAKS IWTGTLIALV IYVLWQTAIQ GNLPRNEFAP
101 VIAAEGQVSV LIETLSKFAQ TGNMDKILSL FSYMAIATSF LGVTLGLFDY
151 IADIFKWNDS VSGRTKTAAL TFLPPLISCL LFPTGFVTAI GYVGLAATVW
201 TGIIPAMLLY RSRKKFGAGK TYKVYGGWL MVWVFLFGIV NIAAQVLSQM
251 ELVPVFKG*

```

m107/a107 94.8% identity in 154 aa overlap

	10	20	30	40	50	60
m107.pep	MVLTFIWAAGGLIADAKPSVLFDTQAPAGTNYWIYAXTALPVCLASFGFHGNVSSLLKYF					
	: : : : :					
a107	MVLTFIWATGGLIADAKLPVLFDTQAPTGTNYWIYVATALPVCLASFGFHGNVSSLLKYF					
	10	20	30	40	50	60

319

	70	80	90	100	110	120
m107.pep	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
a107	KGDAPKVAKSIWTGTLIALVIYVLWQTAIQGNLPRNEFAPVIAAEGQVSVLIETLSKFAQ					
	70	80	90	100	110	120
	130	140	150	160	170	
m107.pep	TGNMDKILSLFSYMAIATSFLGVTLGLFDYIAHLQMERQHLRAAPKPPRX					
a107	TGNMDKILSLFSYMAIATSFLGVTLGLFDYIADIFKWNDSVSGRKTAAALTFPLISCL					
	130	140	150	160	170	180
a107	LFPTGFVTAIGYVGLAATVWTGIIPAMLLYRSRKKFAGAKTYKVYGGGLWMVWVFLFGIV					
	190	200	210	220	230	240

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 381>:

```

g108.seq
1  ATGttgccgg gCTTCAACCG GATATTCAaa cggTTTGCTC CAACACTCGG
51  AACggCGCAT AAAACGCCgc ccTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATCCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTGTCTAT TTGCGCGCG GCAATCTTAC TCGGCGGGTG
201 CGCCGCCGGC GGCAACACAT TCGGCAGCTT AGACGGCGGC ACGGGTATGG
251 GTGGCAGCAT CGTCAAAATG ACGGTAGAAA gccAATGCCG TCGGGAATTG
301 GACAGGCGCA GCGAATGGCG TTGACCGCG CTGGCGATGA GTGCCGAAAA
351 ACAGGCGGAA TGGGAAAACA AGATTGCGG CTGCGCTACC GAAGAAGCAC
401 CTAACCACT GACCGGCAAC GATGTGATGC AGATGCTGaa ccagtccacG
451 CGCaatcagg cacTtgccgc CCtgaccgTC AAAacgggTT CgcctgctT
501 CAaacgcctg tACCGCTaa

```

This corresponds to the amino acid sequence <SEQ ID 382; ORF 108.ng>:

```

g108.pep
1  MLPGFNRIK RFAPTLGTAH KTPPFALSRT GRLIRSYRHK RRGFNKRGIE
51  MNKTLNILPA AILLGGCAAG GNTFGSLDGG TGMGGSIVKM TVESQCRAEL
101 DRRSEWRLTA LAMSAEKQAE WENKICGCAT EEAPNQLTGN DVMQMLNQST
151 RNQALALTV KTVSACFKRL YR*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 383>:

```

m108.seq
1  ATGTTGCCGG GCTTCAACCG GATATTCAAA CGGTTTGTTT CAACACTCGG
51  AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTGTCTAT TTGCGCGTG GCAATCTTAC TCGGCGGCTG
201 CGCCGCCGGA GGCGGTAAAC CATTCGGCAG CTTAGACGGT GGCACAGGCA
251 TGGGCGGCAG CATCGTCAAA ATGGCGGTTG GGAGCCAATG CCGTGCAGAA
301 TTGGACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351 AAAACAGGCG GAGTGGGAAA ACAAGATTG CGCTTGCGTC GCCCAAGAAG
401 CACCCGAACG GATGACCGGC AACGATGTGA TGCAGATGCT GGCTCCGTCC
451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
501 CTTCAAACAC CTGTACCGCT AA

```

This corresponds to the amino acid sequence <SEQ ID 384; ORF 108>:

```

m108.pep
1  MLPGFNRIK RFVPTLGAH KTPPFALSRT GRLIRFYRHK RRGFNKRGIE
51  MNKTLNILPV AILLGGCAAG GGNTFGSLDG GTMGGSIVK MAVGSQCRAE
101 LDKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPERMTG NDVMQMLAPS
151 TRNQALALT AKTVSACFKH LYR*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 108 shows 89.6% identity over a 173 aa overlap with a predicted ORF (ORF 108.ng) from *N. gonorrhoeae*:

m108/g108

320

	10	20	30	40	50	60
m108.pep	MLPGFNRIFKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKLSILPV					
g108	MLPGFNRIFKRFAPTLGTAKHTPPFALSRTGRLIRSYRHKRRGFNRKGIEMNKLSILPA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m108.pep	AILLGGCAAGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA					
g108	AILLGGCAAGG-NTFGSLDGGTGMGGSIVKMTVESQCRAELDRRSEWRLTALAMSAEKQA					
	70	80	90	100	110	
	130	140	150	160	170	
m108.pep	EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAAALTAKTVSACFKHLYRX					
g108	EWENKICGCATEEAPNQLTGNDVMQMLNQSTRNQALAAALTAKTVSACFKRLYRX					
	120	130	140	150	160	170

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 385>:

a108.seq

```

1 ATGTTGCCGG GCTTCAACCG GATATTCAA CGGTTTGTTT CAACACTCGG
51 AACGGCGCAT AAAACGCCGC CCTTCGCGTT ATCCCGAACG GGGCGGCTAA
101 TCAGATTCTA TCGCCATAAA AGGCGGGGTT TCAACCGAAA AGGAATTGAG
151 ATGAATAAAA CCTTGTCTAT TTTGCCGGTG GCAATCTTAC TCGGCGGCTG
201 CGCCGCCGGG GCGGTAACA CATTGCGCAG CTTAGACGGC GGCACAGGTA
251 TGGGCGGCAG CATCGTCAA ATGGCGGTAG AAAGCCAATG CCGTGC GGAA
301 TTGAACAAAC GCAGCGAATG GCGTTTGACC GCGCTGGCGA TGAGTGCCGA
351 AAAACAGGCG GAATGGGAAA ACAAGATTG CGCTTGCGTC GCCCAAGAA
401 CACCCAACCA GCTGACCGGC AACGATGTA TGCAGATGCT GGATCCGTC
451 ACGCGCAATC AGGCACTTGC CGCCCTGACC GCCAAAACGG TTTCCGCCTG
501 CTCAAACAC CTGTACCGCT AA

```

This corresponds to the amino acid sequence <SEQ ID 386; ORF 108.a>:

a108.pep

```

1 MLPGFNRIFK RFVPTLGTAKHTPPFALSRT GRLIRFYRHK RRGFNRRKIE
51 MNKTLNILPV AILLGGCAAG GGNTFGSLDG GTGMGGSIVK MAVESQCRAE
101 LNKRSEWRLT ALAMSAEKQA EWENKICACV AQEAPNQLTG NDVMQMLDPS
151 TRNQALAAALT AKTVSACFKH LYR*

```

m108/a108 96.5% identity in 173 aa overlap

	10	20	30	40	50	60
m108.pep	MLPGFNRIFKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKLSILPV					
a108	MLPGFNRIFKRFVPTLGTAKHTPPFALSRTGRLIRFYRHKRRGFNRKGIEMNKLSILPV					
	10	20	30	40	50	60
	70	80	90	100	110	120
m108.pep	AILLGGCAAGGNTFGSLDGGTGMGGSIVKMAVGSQCRAELDKRSEWRLTALAMSAEKQA					
a108	AILLGGCAAGGNTFGSLDGGTGMGGSIVKMAVESQCRAELNKRSEWRLTALAMSAEKQA					
	70	80	90	100	110	120
	130	140	150	160	170	
m108.pep	EWENKICACVAQEAPERMTGNDVMQMLAPSTRNQALAAALTAKTVSACFKHLYRX					
a108	EWENKICACVAQEAPNQLTGNDVMQMLDPSTRNQALAAALTAKTVSACFKHLYRX					
	130	140	150	160	170	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 387>:

This corresponds to the amino acid sequence <SEQ ID 388; ORF 109.ng>:

g109.pcp

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 389>:

m109.seq

This corresponds to the amino acid sequence <SEQ ID 4; ORF 109>:

m109.pgp

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 109 shows 92.9% identity over a 126 aa overlap with a predicted ORF (ORF 109.ng) from *N. gonorrhoeae*:

m109/g109

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 391>:

a109.seq

1 ATGTATTATC GCCGGGTGT GGGGCTATCC GATGGACTTG GCGATTTGGC
51 AGCCGGTATT GAGCGTAGCC TTGTGCTAG GCGTATACTT ACCGCTTTTG
101 GAAGCGGGCA TGGAATGAC GCGCAAAGGC AAACCAACCC AATCCGCCG
151 CACCGTGGTG TTCTCTCCG CTTGGTCAAT CCGGTTTTCG GCTGGGCGTT

a109.pep

1 MYYRRVVGLS DGLGLDAAGI ERLSLGRRRII TAFSGSHGND AQRQNHPIRR
51 HRGVLFRLLVN PVFGWALTML LDNLGLIGCK ERSAQLGFTG RVLIPVVGFL
101 ILCVAMGAVG MLPGIPPFLE HFKSLG*

		10	20	30	40	50	60
m109.pep		MYRRVRVGLSDGLGDLAAGIERSLGRRRILTAFGSGHGND	AQRQNHP	IRRH	RGVLF	RLVN	
		:					
a109		MYRRRVVGLSDGLGDLAAGIERSLGRRRILTAFGSGHGND	AQRQNHP	IRRH	RGVLF	RLVN	
		10	20	30	40	50	60
		70	80	90	100	110	120
m109.pep		PVFGWALTMLLDNLGLIGCKERSAQLGFAGRVLI	PAVGFLIL	CVAMGAV	GMLPGI	PPFLE	
				:	:		
a109		PVFGWALTMLLDNLGLIGCKERSAQLGFTGRVLI	PVVGFLIL	CVAMGAV	GMLPGI	PPFLE	
		70	80	90	100	110	120
m109.pep		HF	KSLGX				
a109		HF	KSLGX				

```

g111.seq
1  ATGCCGCTCTG  AAACACGCCT  GCCGAACCTT  ATCCGCGCCT  TGATATTTGC
51  CCTGGGTTTC  ATCTTCTCTG  ACGCGTGTTC  GGAacaaacC  GCGCAaaccg
101  TTACCTGCA  AGCGCAAAAG  ATGGGTACGA  CctATACCGT  CAAATACCTT
151  TCAAATAATC  GGGACAAACT  CCGCTCCCCT  GCCAAAATAC  AAAAGCGCAT
201  TGATGATGCG  CTTAAAGAAG  TCAACCGGCA  GATGtccaCC  TACCAGACCG
251  ATTCCGAAAT  CAGCCGGTt  atcacagan  atgctggaga  gctcttcgcy
301  ntntcatgcag  ntctataac  tgattccgcc  gaagactgtc  tgcctaatac
351  qcctatctca  tcqqcgcctc  qa

```

g111.pep

1	MPSETRLPNL	IRALIFALGF	<u>IFLNACSEQT</u>	AQTVTLQGET	MGTTYTVKYL
51	SNNRDKLPSP	AKIQKRIDDA	LKEVNRQMST	YQTDSEISRF	IQTAGELFAH
101	ASITDSAEDC	LPNTPISSAL	*		

```
m111.seq
1  ATGCCGCTCTG  AAACACGCCT  GCCGAAC TTT  ATCCGCGTCT  TGATATTGTC
51  CCTGGGTTTC  ATCTTCCTGA  ACGCGTGT  C  GGAACAAACC  GCGCAAAACG
101  TACCCCTGCA  AGGCGAAACG  ATGGGCAC  GA  CTTATAYCGT  CAAATACCTT
151  TCAAATAATC  GGGACAAACT  CCCCTCAC  CT  GCCGAAATAC  AwAAACGCAT
201  CGATGACGCG  CTTAAAGAAk  TCAACCGy  A  GATGTCCACC  TATCAGCCCG
251  ACTCCGAAAT  CAGCCGGTTC  AACCAACA  C  CAGCCGCGAA  GCCCTCTCCG
301  ATTTCAAGCG  ACTTCGCACA  CGTACTGC  C  GAAGCCGTCC  GCCTGAACCG
351  CCTGACACAC  GGCGCGCTGG  ACGTAACCT  G  CGGCCCTTGT  TGCAACCTTT
401  GGGGATTCGG  CCCCGACAAA  TCCGTTACC  C  GTGAACCGTC  GCCGGAACAA
451  ATCAAAACAG  CGGCATCTTA  TAGGGGCA  T  GACAAAAATCA  TTTTGAAACA
501  AGGCAAGAT  TACGCTTCCT  TGAGCAAA  C  CCACCCCAAG  GCCTATTTTG
551  ATTTATCTTC  GATTGCCAAA  GGCTTCGG  C  TTGATAAAGT  TCGCGGCGAA
```


323

```

601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCGTGGCGC ATCGGTATCG
701 AGCAGCCCAA TATCGTCCAA GCGGCAATA CGCAGATTAT CGTCCCCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATT TCCACGTCGA
801 TAAAAACGGC AAACGCCTCT CCCATATCAT CAACCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCATG
901 ACGGCGGACG GCTTGTCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAATCTCG TGTTCCTCTG ATTGTCAGGG
1001 ATAAAGCGCG cTACCGCACC GCCATGTCTT CCGAATTGTA AAAACTGcTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 396; ORF 111>:

```

m111.pep
  1 MPSETRLPNF IRVLIFALGF IFLNACSEQT AQTVTLQGET MGTYYXVKYL
 51 SNNRDKLPSP AEIXKRIDDA LKEXNRXMS YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVRLNRLTH GALDVTVGPL VNLWGFPGDK SVTREPSPEQ
151 IKQAASYTGI DKIIILKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAEG
201 LEKYGIQNYL VEIGGELHKG GKNARGEPRW IGIEQPNIVQ GGNTQIIVPL
251 NNRSLATSGD YRIFHVDKNG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

ORF 111 shows 88.7% identity over a 97 aa overlap with a predicted ORF (ORF 111.ng) from *N. gonorrhoeae*:

m111.pep/g111.pep

	10	20	30	40	50	60
m111.pep	MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTYYXVKYLSNNRDKLPSP					
g111	MPSETRLPNLIRALIFALGFIFLNACSEQTAQTVTLQGETMGTYYTVKYLNNRDKLPSP					
	10	20	30	40	50	60
m111.pep	AEIXKRIDDALKEKNRXYMSYQPDSEISRFNQHTAGKPLRISDFAHVTA EAVRLNRLTH					
g111	AKIQKRIDDALKEVNRQMSYQTDSEISRFIQTXAGELFAXHAXSITDSAEDECLPNTPI					
	70	80	90	100	110	120
m111.pep	AEIXKRIDDALKEKNRXYMSYQPDSEISRFNQHTAGKPLRISDFAHVTA EAVRLNRLTH					
g111	AKIQKRIDDALKEVNRQMSYQTDSEISRFIQTXAGELFAXHAXSITDSAEDECLPNTPI					
	70	80	90	100	110	120
m111.pep	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASYTGIDKIIILKQKDYASLSKTHPK					
g111	SALX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 397>:

a111.seq

```

  1 ATGCCGTCTG AAACACGCCT GCCGAACCTT ATCCGCACCT TGATATTGTC
 51 CCTGAGTTTT ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAAACG
101 TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
151 TCAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
201 CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCCTCCGC
301 ATTTCAAGCG ACTTCGCACA CGTACTGCC GAAGCCGTCC ACCTGAACCG
351 CCTGACACAC GCGCGCTGG ACGTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCGG CCCGACAAA TCCGTACCC GTGAACCGTC GCCGGAACAA
451 ATCAACAAG CAGCATCTTA TACGGGCATA GACAAATCA TTTGAAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTGCG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TGCGGGCGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGGGCG ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GCGGCAATA CGCAGATTAT CGTCCCCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATT TCCACGTCGA

```

324

801 TAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
 851 CCATCAGCCA CAACCTCGCC TCCATCAGCG TGGTCGCAGA CAGTGCGATG
 901 ACGGCGGACG GCTTGTCCAC AGGATTATTC GTATTGGGCG AAACCGAAGC
 951 CTTAAAGCTG GCAGAGCGCG AAAAATCGC TGTTCCTG ATTGTCAGGG
 1001 ATAAAGCGCG CTACCGCACC GCCATGTCTT CCGAATTGA AAAACTGCTC
 1051 CGCTAA

This corresponds to the amino acid sequence <SEQ ID 398; ORF 111.a>:

a111.pep

1 MPSETRLPNF IRTLIFALSF IFLNACSEQT AQTVTLQGET MGTITYTVKYL
 51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMSY YQPDSEISRF NQHTAGKPLR
 101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFPGDK SVTREPSPEQ
 151 IKQAASVTGI DKIIILKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVGAE
 201 LEKYGIQNYL VEIGGELHKG GKNARGEPRW IGIEQPNIVQ GGNTQIIIVPL
 251 NNRLATSGD YRIFHVDKSG KRLSHIINPN NKRPIHNLAS SISVADSAM
 301 TADGLSTGLF VLGETEALKL AEREKLVFL IVRDKGGYRT AMSSEFEKLL
 351 R*

m111/a111 97.7% identity in 351 aa overlap

m111.pep	10	20	30	40	50	60
	MPSETRLPNFIRVLIFALGFIFLNACSEQTAQTVTLQGETMGTITYXVKYLSNNRDKLPSP					
a111	MPSETRLPNFIRTLIFALSFIFLNACSEQTAQTVTLQGETMGTITYTVKYLSNNRDKLPSP					
	10	20	30	40	50	60
m111.pep	70	80	90	100	110	120
	AEIXKRIDDALKEXNRXMSYQPDSEISRFNQHTAGKPLRISDFAHVTA EAVRLNRLTH					
a111	AEIQKRIDDALKEVNRQMSYQPDSEISRFNQHTAGKPLRISDFAHVTA EAVHLNRLTH					
	70	80	90	100	110	120
m111.pep	130	140	150	160	170	180
	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASVTGIDKIIILKQKDYASLSKTHPK					
a111	GALDVTVGPLVNLWGFPGDKSVTREPSPEQIKQAASVTGIDKIIILKQKDYASLSKTHPK					
	130	140	150	160	170	180
m111.pep	190	200	210	220	230	240
	AYLDLSSIAKGFVDKVGAELEKYGIQNYLVEIGGELHKGKKNARGEPRWIGIEQPNIVQ					
a111	AYLDLSSIAKGFVDKVGAELEKYGIQNYLVEIGGELHKGKKNARGEPRWIGIEQPNIVQ					
	190	200	210	220	230	240
m111.pep	250	260	270	280	290	300
	GGNTQIIIVPLNNRLATSGDYRIFHVDKNGKRLSHIINPNKRPIHNLASISVADSAM					
a111	GGNTQIIIVPLNNRLATSGDYRIFHVDKSGKRLSHIINPNKRPIHNLASISVADSAM					
	250	260	270	280	290	300
m111.pep	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLVFLIVRDKGGYRTAMSSEFEKLLRX					
a111	TADGLSTGLFVLGETEALKLAEREKLVFLIVRDKGGYRTAMSSEFEKLLRX					
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 399>:

g111-1.seq

1 ATGCCGTCTG AACACGCCT GCCGAACCTT ATCCGCGCCT TGATATTTGC
 51 CCTGGGTTTC ATCTTCCTGA ACGCCTGTTC GGAACAAACC GCGCAAacCG
 101 TTACCCTGCA AGGCGAAACG ATGGGTACGA CCTATACCGT CAAATACCTT
 151 TCAAATAATC GGGACAAACT CCCCTCCCTT GCCAAAATAC AAAAGCGCAT
 201 TGATGATGCG CTTAAAGAAG TCAACCGCA GATGTCCACC TACCAGACCG

g111-1.pap

m111-1.seq

m111-1.pgp

ml11-1/g111-1 96.6% identity in 351 aa overlap

mll1-1.pep MPSETRLNPFIRVLIFALGFIFLNACSEQTATQVTVLQGETMGTTYTVKVLSNNRDKLPSP
| | | | | : | : | : | : | : | : | : | : | : | : | : | :
ql11-1 MPSETRLNPLIRALIFALGFIFLNACSEQTATQVTVLQGETMGTTYTVKVLSNNRDKLPSP

326

	10	20	30	40	50	60
m111-1.pep	70	80	90	100	110	120
	AEIQKRIDDALKEVNRQMSTYQPDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH					
g111-1	AKIQKRIDDALKEVNRQMSTYQTDSEISRFNQHTAGKPLRISSDFAHVTAEAVRLNRLTH					
	70	80	90	100	110	120
m111-1.pep	130	140	150	160	170	180
	GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIIKQKGKDYASLSKTHPK					
g111-1	GALDVTVGPLVNLWGFGPDKSVTREPSPEQIKQAASYTGIDKIIKQKGKDYASLSKTHPK					
	130	140	150	160	170	180
m111-1.pep	190	200	210	220	230	240
	AYLDLSSIAGFGVDKVGAGELEKYGIQNYLVEIGGELHGKGNARGEPWRIGIEQPNIVQ					
g111-1	AYLDLSSIAGFGVDKVGAGELEKYGIQNYLVEIGGELHGKGNARGEPWRIGIEQPNIIQ					
	190	200	210	220	230	240
m111-1.pep	250	260	270	280	290	300
	GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNKRPISHNLASISVVADSAM					
g111-1	GGNTQIIVPLNNRSLATSGDYRIFHVDKNGKRLSHIINPNKRPISHNLASISVVSAMS					
	250	260	270	280	290	300
m111-1.pep	310	320	330	340	350	
	TADGLSTGLFVLGETEALKLAEREKLAFLIVRDKGGYRTAMSSEFEKLLRX					
g111-1	TADGLSTGLFVLGETEALRLAEQEKLAFLIVRDKGGYRTAMSSEFAKLLRX					
	310	320	330	340	350	

g111-1/p44550

sp|P44550|YOJL HAEIN HYPOTHETICAL LIPOPROTEIN HI0172 PRECURSOR >gi|1074292|pir||C64144
 hypothetical protein HI0172 - Haemophilus influenzae (strain Rd KW20) >gi|1573128 (U32702)
 lipoprotein, putative [Haemophilus influenzae Rd] Length = 346
 Score = 349 bits (885), Expect = 2e-95
 Identities = 177/328 (53%), Positives = 240/328 (72%), Gaps = 4/328 (1%)

Query: 23 LNACSEQTAQTVTLQGETMGTTYXVKYLSNNRDKLPSPAEIXKRIDDALKEKNRNMSTYQ 82
 L AC ++T + ++L G+TMGTTY VKYL + S + + I+ LK+ N MSTY+
 Sbjct: 17 LAACQKET-KVISLSGKTMGTTYHVKYLDGGSITATS-EKTHEEIEAILKDVNAKMSTYK 74

Query: 83 PDSEISRFNQHT-AGKPLRISSDFAHVTAEAVRLNRLTHGALDVTVGPLVNLWGFGPDKS 141
 DSE+SRFNQ+T P+ IS+DFA V AEA+RLN++T GALDVTVG+VNLWGFGP+K
 Sbjct: 75 KDSELSRFNQNTQVNTPIEISADFVLAELRLNKVTEGALDVTVGPPVNLWGFGPEKR 134

Query: 142 VTREPSPEQIKQAASYTGIDKIIKQKGKDYASLSKTHPKAYLDLSSIAGFGVDKVGAGEL 201
 ++P+PEQ+ + ++ GIDKI L K+ A+LSK P+ Y+DLSSIAGFGVD+VA +L
 Sbjct: 135 PEKQPTPEQLAERQAWVGIDKITLDTNKEKATLSKALPQVYVDLSSIAGFGVDQVAEKL 194

Query: 202 EKYGIQNYLVEIGGELHGKGNARGEPWRIGIEQPNIVQGGNTQIIVPLNNRSLATSGDY 261
 E+ QNY+VEIGGE+ KGKN G+PW+I IE+P + ++ LNN +A+SGDY
 Sbjct: 195 EQLNAQNYMVEIGGEIRAKGNIEGKPWQIAIEKPTTGERAVEAVIGLNNMGMASSGDY 254

Query: 262 RIFHVDKNGKRLSHIINPNKRPISHNLASISVVADSAMTADGLSTGLFVLGETEALKLA 321
 RI+ ++NGKR +H I+P PI H+LASI+V+A +MTADGLSTGLFVLGE +AL++A
 Sbjct: 255 RIY-FEENGKRFAHEIDPKTGYPHQHLASITVLAPTSMTADGLSTGLFVLGEDKALEVA 313

Query: 322 EREKLAFLIVRDKGGYRTAMSSEFEKL 349
 E+ LAV+LI+R G+ T SS F+KL
 Sbjct: 314 EKNNLAVYLIIRTNGFVTKSSSAFKKL 341

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 403>:

a111-1.seq
 1 ATGCCGCTCTG AACACGCCT GCCGAACTTT ATCCGCACCT TGATATTTGC
 51 CCTGAGTTTT ATCTTCCTGA ACGCCTGTTT GGAACAAACC GCGCAAACCG
 101 TTACCCTGCA AGGTGAAACG ATGGGCACGA CCTATACCGT CAAATACCTT
 151 TCAAATAATC GGGACAAACT CCCCTCACCT GCCGAAATAC AAAAGCGCAT
 201 CGATGACGCG CTTAAAGAAG TCAACCGGCA GATGTCCACC TATCAGCCCG
 251 ACTCCGAAAT CAGCCGGTTC AACCAACACA CAGCCGGCAA GCCCTCCCGC
 301 ATTTCAAGCG ACTTCGCACA CGTTACTGCC GAAGCCGTCC ACCTGAACCG

327

```

351 CCTGACACAC GGC GCGCTGG ACCTAACCGT CGGCCCTTG GTCAACCTTT
401 GGGGATTCCG CCCCGACAAA TCCGTTACCC GTGAACCGTC GCCGGAACAA
451 ATCAAACAAG CAGCATCTTA TACGGGCATA GACAAATCA TTTTGAAACA
501 AGGCAAAGAT TACGCTTCCT TGAGCAAAAC CCACCCCAAG GCCTATTGG
551 ATTTATCTTC GATTGCCAAA GGCTTCGGCG TTGATAAAGT TCGGGCGGAA
601 CTGGAAAAAT ACGGCATTCA AAATTATCTG GTCGAAATCG GCGGCGAGTT
651 GCACGGCAAA GGCAAAAACG CGCGCGGCGA ACCTTGCGCG ATCGGCATCG
701 AACAGCCCAA CATCGTCCAA GGC GGCAATA CGCAGATTAT CGTCCCGCTG
751 AACAAACGTT CGCTTGCCAC TTCCGGCGAT TACCGTATTT TCCACGTCGA
801 TAAAGCGGC AAACGCCTCT CCCATATCAT TAATCCGAAC AACAAACGAC
851 CCATCAGCCA CAACCTCGCG TCCATCAGCG TGGTCGCGA CAGTGCAGTG
901 ACGGCGGACG GCTTGTCAC AGGATTATTC GTATTGGCG AAACCGAAGC
951 CTTAAAGCTG GCAGAGCGCG AAAAAGCTCG TGTTCCTG ATTGTCAGGG
1001 ATAAAGCGG CTACCGCAC GCCATGTCT CCGAATTGA AAAAGTCTC
1051 CGCTAA

```

This corresponds to the amino acid sequence <SEQ ID 404; ORF 111-1.a>:

a111-1.pep

```

1 MPSETRLNPF IRTLIFALSF IFLNACSEQT AQTVTLOGET MGTITYTVKYL
51 SNNRDKLPSP AEIQKRIDDA LKEVNRQMST YQPDSEISRF NQHTAGKPLR
101 ISSDFAHVTA EAVHLNRLTH GALDVTVGPL VNLWGFDPK SVTREPSPEQ
151 IKQAASYTGI DKILKQKGD YASLSKTHPK AYLDLSSIAK GFGVDKVAGE
201 LEKYGIQNYL VEIGGELHGK GKNARGEPRW IGIEQPNIVQ GGNTQIIIVPL
251 NNRLATSGD YRIFHVDKSG KRLSHIINPN NKRPISHNLA SISVVADSAM
301 TADGLSTGLF VLGETEALKL AEREKLAVFL IVRDKGGYRT AMSSEFEKLL
351 R*

```

a111-1/m111-1 98.9% identity in 351 aa overlap

	10	20	30	40	50	60
a111-1.pep	MPSETRLNPFIRTLIFALSFIFLNACSEQT	AQTVTLOGETMGTITYTVKYL	SNNRDKLPSP			
m111-1	MPSETRLNPFIRVLIFALGFIFLNACSEQT	AQTVTLOGETMGTITYTVKYL	SNNRDKLPSP			
	10	20	30	40	50	60
	70	80	90	100	110	120
a111-1.pep	AEIQKRIDDALKEVNRQMSTYQPDSEISRF	NQHTAGKPLRISSDFAHVTA	EAVHLNRLTH			
m111-1	AEIQKRIDDALKEVNRQMSTYQPDSEISRF	NQHTAGKPLRISSDFAHVTA	EAVHLNRLTH			
	70	80	90	100	110	120
	130	140	150	160	170	180
a111-1.pep	GALDVTVGPLVNLWGFDPKSVTREPSPEQ	IKQAASYTGIDKILKQKGDYASLSKTHPK				
m111-1	GALDVTVGPLVNLWGFDPKSVTREPSPEQ	IKQAASYTGIDKILKQKGDYASLSKTHPK				
	130	140	150	160	170	180
	190	200	210	220	230	240
a111-1.pep	AYLDLSSIAKGFVDKVAGELEKYGIQNYL	VEIGGELHGKGNARGEPRWIGIEQPNIVQ				
m111-1	AYLDLSSIAKGFVDKVAGELEKYGIQNYL	VEIGGELHGKGNARGEPRWIGIEQPNIVQ				
	190	200	210	220	230	240
	250	260	270	280	290	300
a111-1.pep	GGNTQIIIVPLNNRLATSGDYRIFHVDKSG	KRLSHIINPNNKRPISHNLA	SISVVADSAM			
m111-1	GGNTQIIIVPLNNRLATSGDYRIFHVDKSG	KRLSHIINPNNKRPISHNLA	SISVVADSAM			
	250	260	270	280	290	300
	310	320	330	340	350	
a111-1.pep	TADGLSTGLFVLGETEALKLAEREKLAVFL	IVRDKGGYRTAMSSEFEKLLRX				
m111-1	TADGLSTGLFVLGETEALKLAEREKLAVFL	IVRDKGGYRTAMSSEFEKLLRX				
	310	320	330	340	350	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 405>:

g114.seq

```

1 ATGGCTTCCA TCACTTCGCC GCTGCACGGG GCGCAGCAGG AATGCAGCAA
51 GACTTTTTTA TGTCCGCCGG GCGGACGAG TATGGGGCGG TCAATGTCGG

```

1 MASITSPH LG AQECSKTF L CPPGGTSMGR SMSVTVGLFC VSINLTISVE
51 YGQSGYFTR AECKTGCQGI SPSC LNERIV CEVTIKWSSS ETSTSDMACA
101 SRLVNMSSC EGSGEPPGWL CAIRLSAYS SNASLTTRSM *

1	ATGGCTTCCA	TCACCTCGCC	GCTGCACGGG	GCGCACAGAG	AATGCAGCAA
51	GACTTTTTTA	TGTCCACCGG	GCGGGACGAG	TATAGGGCGG	TCAATGTCGG
101	TAACGGTAGG	TTTGTTTTGT	GTTCACATA	ACTTAACAAT	ATCTGTTGAA
151	TACGGTTGAA	CGGGCTATT	TATCAGACCC	CGCGCATGTA	AAACAGAGTG
201	TACGGGCATG	AACCCGAGCT	GTCGTAACGA	ACAGACGCTT	TGCGAKGTAA
251	CGATAAAATG	GTCGAGCAGC	GACACATCGA	CCAGCGACAT	TGCCTGTGCC
301	AGCCGCGCTT	TGAACATATG	GTCTTCCTGC	GAArGTTcG	GCGAGCCGcC
351	CgGatGGTTG	TGCGCAATAA	TcAGGCTGTC	GGCATATTcG	TCCAATGCCA
401	GTTTGACGAT	TTTCGCGGATG	TAA		

1 MASITSPLHG AHRECSKFTL CPPGGTSIGR SM SVTVGLFC VSINLTISVE
51 YGXSGYFIRA AACKTECQGI NPSCLEQTL CKVTIKWSSS DTSTSDIACA
101 SRLVNMSSC EXSGEPPGWL CAIIRLSAYS SNASLTISRM *

	10	20	30	40	50	60
m114.pep	MASITSP	LHG	AHRECS	KTF	CP	PGGTSIGR
g114	MASITSP	LHG	AQEQCS	KTF	CP	PGGTSMGR
	10	20	30	40	50	60
	70	80	90	100	110	120
m114.pep	AACKTEC	QGINP	SCLNEQ	TLCXV	TIKWSS	SDTSTSD
g114	AECKTGC	QGISP	SCLNERT	VCEVTI	KWSSSE	TSTSDMA
	70	80	90	100	110	120
	130	140				
m114.pep	CAIIRLS	AYSSN	ASLTIS	RMX		
g114	CAIIRLS	AYSSN	ASLTIS	RMX		
	130	140				

1	ATGCCGGAGG	CAAGCATCGC	CTCCATCACT	TCGCCGCTGC	ACGGGGCGCA
51	ACAGGAATGC	AGCAAGACTT	TTTTATGTCC	GCCGGGCGGG	ACGAGTATGG
101	GGCGGTCAAT	GTCGGTAAAC	GTAGGTTTGT	TTTGTGTTTC	CATTAACTTA
151	ACGATATCTG	TCGAATACGG	TTGAAGCGGC	TATTTTATCA	GAGCCGCCGC
201	ATGTAAACA	GGGTGTCAGG	CGTACAGCCC	GAGCTGCGCT	AACGAACGGA
251	CGGTTTGC	CGTTACGATA	AAATGGTCGA	GCAGCGACAC	ATCGACCAGC
301	GACATTGCCT	GCGCCAGCCG	CCTTGTGAAC	ATGATGCTT	CCTGCGAAGG
351	TTCCGGCGAG	CCGCCGGATT	GGTTGTGCGC	GATAATCAGG	CTGTCCGCAT
401	ATTCTGCCAA	TGCCAGTTTG	ACAATTCAC	GGATGTAA	

This corresponds to the amino acid sequence <SEQ ID 410; ORF 114.a>:

a114.pep

```

1 MPEASIASIT SPLHGAQDEC SKTFLCPPGG TSMGRSMSVT VGLFCVSINL
51 TISVEYG*SG YFIRAAACKT GCQGISPSCN NERTVCAVTI KWSSSDTSTS
101 DIACASRLVN MMSSCEGSGE PPGWLCAIIR LSAYSSNASL TISM*
```

m114/a114 92.9% identity in 140 aa overlap

```

      10      20      30      40      50
m114.pep  MASITSPLHGAHRECSKTFCLPPGGTSIGRSMSTVGLFCVSINLTISVEYGXSG
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a114      MPEASIASITSPLHGAQDECSTFLCPPGGTSMGRSMSTVGLFCVSINLTISVEYGXSG
      10      20      30      40      50      60

      60      70      80      90     100     110
m114.pep  YFIRAAACKTECQGINPSCLEQTLXVTIKWSSSDTSTSDIACASRLVNMSSCEXSSE
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a114      YFIRAAACKTGCQGISPSCNERTVCAVTIKWSSSDTSTSDIACASRLVNMSSCEGSSE
      70      80      90     100     110     120

      120     130     140
m114.pep  PPGWLCAIIRLSAYSSNASLTISM*
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a114      PPGWLCAIIRLSAYSSNASLTISM*
      130     140
```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 411>:

g117.seq

```

1 atggtcgacg aactcgacCT GCTGCCGAT GCCGTCGCCG CCACCCTGCT
51 TGCCGACATC GGACGCTACG TCCCCGATTG GAACCTATTG GTTCCGAGC
101 GCTGCAACAG CACCGTCGCC GAGCTGGTCA AAGGTgtgga CGAAGTGCAG
151 AAACCTACCC ACTTCGCCCC GGTGGACAGC CTCGCCACGC CGAAGAAGACG
201 CGCACAGCAA GCGGAAACCA TCGGAAAAT GCTGCTGGCg atggttaccg
251 Acatccgct cgtatTAATC AAACCTGGCGA TGCGTaccg caccCTGcta
301 tTTTtaagCA ACGCCCCCGA CAGCCCTGAA AAACgcgccG TCgcaaaAga
351 aacccTCGAC ATCTTCGCCC CGCTCGCCAA CCGCTTGGGC GTGTGGCAGC
401 TCAAAATGGCA GCTCGAAGAT TTGGGCTTCC GCCATCAAGA ACCCGAAAAA
451 TACCGCGAAA TCGCCTGCT TTTGGACGAA AAACGCACCG AACGCCTCGA
501 ATACATCGAA AACTTCCTCG ATATCCTGCG TACGGAATC AAAAAATACA
551 ATATCCACTT TGAAGTCGCC GGCCGTCCGA AACACATCTA CTCCATTTAC
601 AAAAAATAGG TGAAGAAAAA ACTCAGCTTC GACGgccTGT TCGACATCCG
651 CGCCGTGCGG ATTCTGGTCG ATACCGTCCC CGaGTGTTAC ACCACGCTGG
701 gcaTCGTCCA CAGCCTCTGG CAGCCATTC CCGCGgagtt CGAcgactAC
751 ATCGCCAACC CCAAAGgcaA CGgttATAAA AGtTTGCACA CCGTCATCGT
801 cggcccGGAA gacaaagggtg tggaaGtgCA AATCCGCACC TTCGatGC
851 accAATTcaa CgaatTcggT gtcgcccGCC ACTGGCGtta caaagaaggc
901 ggcaaaaggcg attccGCtA cgaacaaAAA ATcgccTggt TGCgcaACT
951 CTTGGACTGG CGCGAAAATA TGGCGAAAAG CGGCAAGGAA GACCTCGCCG
1001 CCGCCTTCAA AACCGAGCTT TTCAACGACA CGATTTATGT TTTGACCCCG
1051 CACGGCAAG TCCTCTCTCT GCCAACGGGC GCAACCCCA TCGACTTCGC
1101 CTACGCCCTG CACAGCAGCA TcggCGACCG CTGCCGGGGC GCGAAAGTCG
1151 AaggGCAGAT TGTGCCGCTG TCCACCCCGC TCGAAAACGG ACAGCGCGTC
1201 GAAATcatta cCGCcaaAGA AGGGCATCCT TCCGTCAACT GGCTTTACGA
1251 AGGctgGGtc aAATCCGGCA AGGCCATCGG caaAATCCGC GCCTAcacCC
1301 GCCAGcaaAa cgCgaCACC GTGCGCGAAG AAGGCCGTGT CCAACTCGAC
1351 AAGCAGCTTG CCAAACTCAC GCCCAAAACC AACCTGCAAG AGCTTgcca
1401 aaATCTCGGC tacaaAAAGC cagaagacct ctacacCGCc gtcggacaag
1451 gcgaaatttc caaccgccc atCaaaaaag cctgcgccac GCTgaacgaa
1501 ccgccccCGG TGCCCGTCAG CGCAACCACC ATCGTCAAC AGTCCAAAAT
```

```

1551 CAAAAAAGGT GGCAAACCG GCGTGCTCAT CGACGGCGAA GACGGCTTGA
1601 TGACCACGCT TGCCAAATGC TGCAAACCG CGCCGCCCGA CGATATTGCC
1651 GGCTTCGTTA CCCGCGAGCG CGGCATTTC GTCCACCGCA AAACCTGCCC
1701 CTCTTCCGA CACCTTGCCG AACACGCGCC CGAAAAAGTA CTGGACGCAA
1751 GTTGGGCGGC GTTGCAGGAA GGGCAAGTGT TCGCCGTCGA TATCGAAATC
1801 CGCGCCCAAG ACCGCTCCGG GCTTTTGCGC GACGTATCCG ACGCGCTCGC
1851 CCGCCACAAA CTCAACGTTA CCGCCGTGCA AACCCAGTCC CGCGACTTGG
1901 AAGCCAGCAT GAGGTTACG CTCGAAGTCA AACAAGTCAA CGacCTCCCG
1951 CGCGTCCTCG CCGGCCTCGG CGATGTCAA GCGGTATTGA GCGTTACCCG
2001 GCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 412; ORF 117.ng>:

g117.pep

```

1 MVDEL DLLPD AVAATLLADI GRYVPDWNLL VSERCNSTVA ELVKGVDEVQ
51 KLTHFARVDS LATPEERAQQ AETMRKMLLA MVTDIRVLI KLAMRTRTLL
101 FLSNAPDSPE KRAVAKETLD IFAPLANRLG VWQLKWQLED LGFRHQEPEK
151 YREIALLLDE KRTERLEYIE NFDILRTEL KKYNIHFEVA GRPKHIYSIY
201 KKMVKKKLSF DGLFDIRAVR ILVDTVPECY TTLGIVHSLW QPIPGFDDY
251 IANPKGNGYK SLHTVIVGPE DKGVEVQIRT FDMHQFNEFG VAAHWRYKEG
301 GKGD SAYEQ IAWLRQLLDW RENMAESGKE DLAAAFKTEL FNDTIYVLTP
351 HGKVLSPPTG ATPIDFAYAL HSSIGDRCRG AKVEGQIVPL STPLENGQRV
401 EII TAKEGHP SVNWLYEGWV KSGKAIGKIR AYIRQQNADT VREEGRVQLD
451 KQLAKLTPKP NLQELAENLG YKKPEDLYTA VGQGEISNRA IQKACGTLNE
501 PPPVPVSATT IVKQSKI KKG GKTGV LIDGE DGLMTTLAKC CKPAPPDDIA
551 GFVTRERGIS VHRKTCPSFR HLAEHAPEKV LDASWAALQE GQVFAVDIEI
601 RAQDRSGLLR DVSDALARHK LNVTA VQTQS RDLEASMRFT LEVKQVNDLP
651 RVLAGLGDVK GVLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 413>:

m117.seq (partial)

```

1 ..GTGAACTCA AGAAATACAA TGTCCATTTC GAAGTCGCCG GCCGCCCGAA
51 ACACATCTAC TCCATTTACA AAAAAATGGT GAAGAAAAAA CTCAGCTTCG
101 ACGGCCTCTT TGACATCCGC GCCGTGCGAA TTCTGGTTGA TACCGTCCCC
151 GAGTGTTACA CCACGCTGGG TATCGTCCAC AGCCTCTGGC AGCCCATTC
201 CGGCGAGTTC GACGACTACA TCGCCAATCC CAAAGGCAAC GGCTATAAAA
251 GTTTGCACAC CGTCATCGTC GGCCCGGAAG ACAAGGCGT GGAAGTACAA
301 ATCCGCACCT TCGATATGCA CCAATCAAC GAATTCGGTG TCGCCGCCCA
351 CTGgCGTTAC AAAGAGGGCG GCAAGGGCGA TTCCGCTTAC GAACAGAAAA
401 TCGCCTGGTT GCGCCAACTC TTGGACTGGC GCGAAACAT GGCGGAAAGC
451 GGCAAGGAAG ACCTCGCCGC CGCCTTCAA ACCGAGCTTT TCAACGACAC
501 GATTTATGTT TTGACCCCGC ACGGCAAAGT CCTCTCCCTG CCCACGGGCG
551 CGACCCCAT CACTTCGCC TACGCCCTGC ACAGCAGCAT CGGCACCGT
601 TGCCGCGGTG CGAAAGTCGA AGGGCAGATT GTGCCGTGT CCACCCCGCT
651 CGAAACGGA CAGCGCTCG AAATCATTAC CGCCAAAGAA GGCATCCTT
701 CCGTCAACTG GCTTTACGAA GGCTGGGTCA AATCCAACAA GGCAATCGGC
751 AAAATCCGCG CCTACATCCG CCAGCAAAAC GCCGACACCG TCGCGGAAGA
801 AGGCCGCGTC CAACTCGACA AACAGCTTGC CAACTCACG CCCAAACCCA
851 ACCTGCAAGA GCTTGCGGAA AATCTCGGCT ACAAAGGCC AGAAGACCTC
901 TACACCGCCG TCGGACAAGG CGAAATTTCC AACCGCGCCA TCCAAAAGC
951 CTGCGGCACg CTGAACGAAC CGCCGCCCGT ACCCGTCAGC GAAACCAACA
1001 TCGTCAAACA GTCCAAAATC AAAAAAGGCG GCAAAAACGG CGTGCTCATC
1051 GACGGCGAAG ACGGTCTGAT GACCACGCTT GCCAAATGCT GCAAACCCGC
1101 GCCGCCGAC GATATTATCG GCTTCGTTAC CCGCGAGCGC GgCATTTCAG
1151 TGCACCGCAA AwyyTkCyCG TCTTTCCAAC ACCTCGCCGA ACACGCGCCC
1201 GAWAAAAGTGC TGGACGCAAG CTGGGCGGCA TTGCAGGAAG GACAAGTATT
1251 CGCCGTCGAT ATCGAAATCC GCGCCCAAGA CCGCTCCGGG CTTTTGCGCG
1301 ACGTATCCGA CGCGCTCGCC CGCCACAAAC TCAACGTTAC CGCCGTGCAA
1351 ACCCAGTCCC GCGACTTGA AGCCAGCATG AGGTTACAGC TCGAAGTCAA
1401 ACAAGTCAAC GACCTCCCGC GCGTCTCCGC CAGCCTCGGC GACGTCAAAG
1451 GCGTATTGAG CGTTACCCGG CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 414; ORF 117>:

m117.pep (partial)


```

1...VKLKKNVHF EVAGRPKHIY SIYKKMVKKK LSFDGLFDIR AVRILVDTV
51 ECYTTLGIVH SLWQIPGFEF DDYIANPKGN GYKSLHTVIV GPEDKGVEVQ
101 IRTFDMHQFN EFGVAHWRY KEGGKGSAY EQKIAWLRQL LDWRENMAES
151 GKEDLAAAFK TELFNDTIYV LTPHGKVLSL PTGATPIDFA YALHSSIGDR
201 CRGAKVEGQI VPLSTPLENG QRVEIITAKE GHPSVNWLYE GWVKSNAIG
251 KIRAYIRQON ADTVREEGRV QLDKQLAKLT PKPNLQELAE NLGYKKPEDL
301 YTAVGQGEIS NRAIQACGT LNEPPVPVVS ETTIVKQSKI KKGKNGVLI
351 DGEDGLMTTL AKCCKPAPPD DIIGFVTRER GISVHRKXXX SFQHLAEHAP
401 XKVLDASWAA LQEGQVFAVD IEIRAQDRSG LLRDVSDALA RHKLNVTAVQ
451 TQSRDLEASM RFTLEVQVN DLPRVLASLG DVKGVLVTR L*

```

Computer analysis of this amino acid sequence gave the following results:

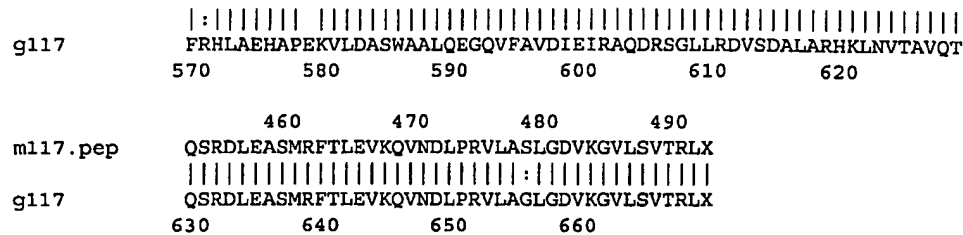
Homology with a predicted ORF from *N. gonorrhoeae*

ORF 117 shows 97.6% identity over a 490 aa overlap with a predicted ORF (ORF 117.ng) from *N. gonorrhoeae*:

m117/g117

m117.pep				10	20	30
				VKLKKNVHF	EVAGRPKHIY	SIYKKMVKKK
g117				:	:	:
	150	160	170	180	190	200
m117.pep		40	50	60	70	80
		SFDGLFDIR	AVRILVDTV	PECYTTLGIVH	SLWQIPGFEF	DDYIANPKGN
g117		SFDGLFDIR	AVRILVDTV	PECYTTLGIVH	SLWQIPGFEF	DDYIANPKGN
	210	220	230	240	250	260
m117.pep		100	110	120	130	140
		PEDKGVEVQ	IRTFDMHQFN	EFGVAHWRY	KEGGKGSAY	EQKIAWLRQL
g117		PEDKGVEVQ	IRTFDMHQFN	EFGVAHWRY	KEGGKGSAY	EQKIAWLRQL
	270	280	290	300	310	320
m117.pep		160	170	180	190	200
		KEDLAAAFK	TELFNDTIYV	LTPHGKVLSL	PTGATPIDFA	YALHSSIGDR
g117		KEDLAAAFK	TELFNDTIYV	LTPHGKVLSL	PTGATPIDFA	YALHSSIGDR
	330	340	350	360	370	380
m117.pep		220	230	240	250	260
		PLSTPLENG	QRVEIITAKE	GHPSVNWLYE	GWVKSNAIG	KIRAYIRQON
g117		PLSTPLENG	QRVEIITAKE	GHPSVNWLYE	GWVKSNAIG	KIRAYIRQON
	390	400	410	420	430	440
m117.pep		280	290	300	310	320
		LDKQLAKLT	PKPNLQELAE	NLGYKKPEDL	YTAVGQGEIS	NRAIQACGT
g117		LDKQLAKLT	PKPNLQELAE	NLGYKKPEDL	YTAVGQGEIS	NRAIQACGT
	450	460	470	480	490	500
m117.pep		340	350	360	370	380
		TTIVKQSKI	KKGKNGVLID	GEDGLMTTL	AKCCKPAPPD	DIIGFVTRER
g117		TTIVKQSKI	KKGKNGVLID	GEDGLMTTL	AKCCKPAPPD	DIIGFVTRER
	510	520	530	540	550	560
m117.pep		400	410	420	430	440
		FQHLAEHAP	XKVLDASWAA	LQEGQVFAVD	IEIRAQDRSG	LLRDVSDALA
g117		FQHLAEHAP	XKVLDASWAA	LQEGQVFAVD	IEIRAQDRSG	LLRDVSDALA

332



The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 415>:

a117.seq

```

1  ATGGTTCATG  AACTCGACCT  GCTCCCGCAT  GCCGTCGCCG  CCACCTGCT
51  TGCCGACATC  GGACGCTACG  TCCCGGACTG  GAACCTATTG  GTTTCGGAAC
101 GCTGCAACAG  TACCGTCGCC  GAGCTGGTCA  AAGGTGTGGA  CGAAGTGCAG
151 AAACACACCC  ACTTCGCCCC  GGTGGACAGC  CTCGCCACGC  CGGAAGAACG
201 CGCCCAGCAG  GCAGAAACTA  TGCAGGAAAT  GCTGCTGGCG  ATGGTTACCG
251 ACATCCGCGT  CGTGTTAATC  AAACGGCGCA  TCGGTACGCG  CACCTGCAA
301 TTTTAAAGCA  ACGCCCCGGA  CAGCCCCGAA  AAACGCGCCG  TCGCCAAAGA
351 AACCTCGAC  ATCTTCGCCC  CGCTCGCCAA  CCGTTTGGGC  GTGTGGCAGC
401 TCAAATGGCA  GCTCGAAGAT  TTGGGCTTCC  GCCATCAAGA  ACCCGAAAAA
451 TACCGCGAAA  TCGCCCTGCT  TTTGGACGAA  AAACGACCCG  AACGCTCGA
501 ATACATCGAA  AACTTCCTTA  ATATCTGCG  TACGGAATC  AAAAAATACA
551 ATATCCACTT  TGAAGTCGCC  GGCCGTCCGA  AACACATCTA  CTCCATTAC
601 AAAAAAATGG  TGAAGAAAAA  ACTCAGCTTC  GACGGGTGT  TCGACATCCG
651 CGCCGTGCGG  ATTCTGGTTG  ATACGCTCCC  CGAGTGTTAC  ACCACACTGG
701 GCATTGTCCA  CAGCCTCTGG  CAGCCATTTC  CCGGCGAGTT  CGACGACTAC
751 ATCGCCAACC  CGAAAGGCAA  CGGCTATAAA  AGTTTGACA  CCGTCATCGT
801 CGGCCCCGAA  GACAAAGGCG  TGGAAGTGCA  AATCCGCACC  TTCGATATGC
851 ACCAATTCAA  CGAATTCGGT  GTCGCCGCGC  ACTGGCGTTA  CAAAGAGGGC
901 GGCAAAGGCG  ATTCCGCCTA  CGAACAAAAA  ATCGCCTGGT  TACGCCAAT
951 TTTGGACTGG  CGCGAAAACA  TGGCGGAAAG  CGGCAAGGAA  GACCTCGCCG
1001 CCGCCTTCAA  AACCAGGCTT  TTCAACGACA  CGATTATGT  TTTGACCCCG
1051 CACGGCAAAG  TCCTCTCCCT  GCCCACAGGC  GCGACCCCA  TCGACTTCGC
1101 CTACGCCCTG  CACAGCAGCA  TCGGCGACCG  TTGCGCGGT  GCGAAAGTCG
1151 AAGGGCAGAT  TGTGCCGTG  TCCACCCCGC  TCGAAAACGG  ACAGCGTGTG
1201 GAAATCATT  CCGCCAAAGA  AGGGCATCCT  TCCGTCAACT  GGCTTTACGA
1251 AGGCTGGGTC  AAATCCAACA  AGGCAATCGG  CAAATCCGC  GCCTACATCC
1301 GCCAGCAAAA  CGCCGACACC  GTGCGCAAG  AAGGCGCGT  CCAACTCGAC
1351 AAACAGCTTG  CCAACTCAC  GCCAAACCC  AACCTGCAAG  AGCTTGCCGA
1401 AAATCTCGGC  TACAAAAAGC  CAGAAGACCT  CTACACCGCC  GTCGACAAAG
1451 GCGAAATTC  CAACCGCGCC  ATCCAAAAG  CCTGCGGCAC  GCTGAACGAA
1501 CCGCCGCCCG  TACCCGTCAG  CGAAACCACC  ATCGTCAAAC  AGTCAAAAT
1551 CAAAAAAGGC  GGCAAAACG  GCGTGCTCAT  CGACGGCGAA  GACGGTCTGA
1601 TGACCAAGCT  TGCCAAATGC  TGCAAAACCG  CGCCGCCCGA  CGACATTGTC
1651 GGCTTCGTTA  CCCGCGATCG  CGGCATTCG  GTACACCGCA  AAACCTGCCC
1701 CTCTTTCCGA  CACCTCGCCG  AACACGCGCC  CGAAAAAGTA  CTGGACGCAA
1751 GTTGGGCGGC  GTTGCAGGAA  GGACAAGTGT  TCGCCGTCGA  TATCGAAATC
1801 CGCGCCCAAG  ACCGCTCCGG  GCTTTGCGC  GACGTATCCG  ACGCGCTCGC
1851 CCGCCACAAA  CTCAACGTTA  CCGCCGTGCA  AACCAGTCC  CGCGACTTGG
1901 AAGCCAGCAT  GAGGTTACAG  CTCGAAGTCA  AACAAGTTAC  CGACCTCCCA
1951 CGCGTCTCG  CCAGCTCCG  CGACGTCAA  GGCGTATTGA  GCGTTACCCG
2001 GCTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 416; ORF 117.a>:

a117.pep

```

1  MVHELDLLPD  AVAATLLADI  GRYVPDWNLL  VSERCNSTVA  ELVKGVDEVO
51  KLTHFARVDS  LATPEERAQQ  AETMRKMLLA  MVTDIRVLI  KLAMRTRTLQ
101  FLSNAPDSPE  KRAVAKETLD  IFAPLANRLG  VWQLKWQLED  LGFRHQEPEK
151  YREIALLLDE  KRTERLEYIE  NFLNILRTEL  KKYNIHFEVA  GRPKHIYSIY
201  KKMVKKKLSF  DGLFDIRAVR  ILVDTVPECY  TTLGIVHSLW  QPIPGEFDDY
251  IANPKNGYK  SLHTVIVGPE  DKGVEVQIRT  FDMHQFNEFG  VAAHWRYKEP
301  GKGDSEYEQK  IAWLRQLLDW  RENMAESGKE  DLAAAFKTEL  FNDTIYVLTG
351  HGKVLSLPTG  ATPIDFAYAL  HSSIGDRCRG  AKVEGQIVPL  STPLENGQRV

```

— 159 —

[illegible]

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 417>:

g117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CCCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
101 AAAACCTCAT CGGTACCGCA TGGTCGCTGG CGCAGGAACA TTATCCTGCC
151 GATGCCGCCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201 GGCAGCAATG GTCGACGAAC TCGACCTGCT GCCCGATGCC GTCGCCGCCA
251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGATTGGAA CCTATTGGTT
301 TCCGAGCGCT GCAACAGCAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351 AGTGCAGAAA CTTACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401 AAGAAGCGCG ACAGCAAGCG GAAACCATGC GGAATGCTCT GCTGGCGATG
451 GTTACCGACA TCCGCGTCGT ATTAATCAAA CTGGCGATGC GTACGCGCAC
501 CCTGCAATTT TTAAGCAACG CCCCAGACAG CCCTGAAAAA CGCGCCGTCG
551 CCAAAGAAAC CCTCGACATC TTCGCCCGCG TCGCCAACCG CTTGGGCGTG
601 TGGCAGCTCA AATGGCAGCT CGAAGATTGG GGCTTCCGCC ATCAAGAACC
651 CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701 GCCTCGAATA CATCGAAAAA TTCCTCGATA TCCTGCGTAC GGAACCAAAA
751 AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC
801 CATTTACAAA AAAATGGTGA AGAAAAACT CAGCTTCGAC GGCCTGTTCG
851 ACATCCGCGC CGTGGCGGAT CTGGTCGATA CCGTCCCGGA GTGTACACCC
901 ACGCTGGGCA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGagttCGA
951 cgactACATC GCCAACCCCA AAGgcaACGg ttATAAAAGt TTGCACACCG
1001 TCATCGTCgg cccGGAagaa aaaggtgtgg aagtgcAAAT CCGCACCTTC
1051 GATATGCacc AATTCaCga ATTCGGTGTC GCCGCCACT GGCCTTACAA
1101 AGAAGGCGGC AAAGGCGATT CCGCTACGA ACAAATAATC GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAAAATATG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCGC CTTCAAAAC CGAGCTTTTC AACGACAGA TTTATGTTTT
1251 GACCCCGCAC GGCAAGTCC TCTCTCTGCC AACGGGCGCA ACCCCCATCG
1301 ACTTCGCTTA CGCCCTGCAC AGCAGCATCG GCGACCGCTG CCGGGGCGCG
1351 AAGTCTGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCCTGCGAAA ATCATTACCG CCAAGAAAGG GCATCCTTCC GTCAACTGGC
1451 TTTAGCAAGG CTGGGTCAA TCCGCAAGG CCATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGTGTCCA
1551 ACTCGACAAG CAGCTTGCCA AACTCAGGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTGC CCGTCAGCGC AACCACCATC GTCAAAACAGT
1751 CCAAATCAA AAAAGGTGGC AAAACCGGCG TGCTCATCGA CGCGGAAGAC
1801 GGCTTGATGA CCACGCTTGC CAAATGTGTC AAACCCGCGC CGCCCGACGA
1851 TATTGCCGCG TTCGTTACCC GCGAGCGCGG CATTTCCGTC CACCGCAAAA
1901 CCTGCCCCCT TTTCCGACAC CTTGCCGAAC ACGCGCCGGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGG CAAGTGTTCT CCGTCGATAT
2001 GGAATATCCG GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAATC AACGTTACCG CCGTGCAAA CAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAA AAGTCAACGA
2151 CCTCCCGCGC GTCTCGCCG GCCTCGGCGA TGTCAAAGGC GTATTGAGCG
2201 TTACCCGGCT TTA

```

This corresponds to the amino acid sequence <SEQ ID 418; ORF 117-1.ng>:

g117-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSCAALP DNDKNLIGTA WSLAQEHYPA
51  DAATPYGEPL PDHFLGAQM VDELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLDILRTELK
251 KYNIHFEVAG RPKHIYSIYK KMKKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKNGYKS LHTVIVGPPE KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTTPH GKVLSLTPGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQVVE IITAKEGHPV VNWLYEGWVK SGKAIGKIRA
501 YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNEP PPVPVSATTI VKQSKIKKGG KTGVLIDGED
601 GLMTTFLAKC KPAPPDDIAG FVTRERGISV HRKTCPSFRH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVNDLPR VLAGLGDVKG VLSVTRL*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 419>:

m117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA

```

```

51  ATTGCGCGAA TGGTTCGACA GCTACTGCGC CGCTCTGCCG GACAACGATA
101 AAAACCTCAT CGGTACCGCA TGGTTGCTGG CGCAGGAACA TTACCCCGCC
151 GATGCCCGCA CGCCGTATGG CGAGCCGCTG CCCGACCACT TCCTCGGCGC
201 GGCACAAATG GTTCATGAAC TCGACCTGCT CCCCGATGCC GTCGCCGCCA
251 CCCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301 TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351 AGTGCAGAAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCCGG
401 AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAATATGCT GCTGGCGATG
451 GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501 CCTGCAATTT TTAAGCAACG CCCCAGCAG CCCCAGAAAA CGCGCCGTCG
551 CCAAGAAAC CCTCGACATC TTCGCCCGC TCGCCAACCG TTTGGGCGTG
601 TGGCAGCTCA AATGGCAGCT CGAAGATTG GGGTTCCGCC ATCAAAAGCC
651 CGAAAAATAC CGCGAAATCG CGCTGCTTTT GGACGAAAAA CGCACCGAAC
701 GCCTCGAATA CATCGAAAC TTCTCAACA TCCTGCGCGG TGAATCAAG
751 AATACAAATG TCCATTTCTG AGTCGCCGGC CGCCCGAAAC ACATCTACTC
801 CATTTACAAA AAAATGGTGA AGAAAAAACT CAGCTTCGAC GGCCTCTTTG
851 ACATCCGCGC CGTGCGAATT CTGGTTGATA CCGTCCCGCA GTGTTACACC
901 ACGCTGGGTA TCGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
951 CGACTACATC GCCAATCCCA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTACAAAT CCGCACCTTC
1051 GATATGCACC AATTCAACGA ATTGCGTGTC GCGGCCCACT GCGGTTACAA
1101 AGAGGGCGGC AAGGGCGATT CCGCTACGA ACAGAAATC GCCTGGTTGC
1151 GCCAACTCTT GGACTGGCGC GAAACATGG CGGAAAGCGG CAAGGAAGAC
1201 CTCGCCCGCG CTTCAAAAC CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAAAGTCC TCTCCCTGCC CACGGGCGCG ACCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCGCTGTCC ACCCGCTCG AAAACGGACA
1401 GCGCGTCGAA ATCATTACCG CCAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGC CTGGGTCAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCGCGTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCAGGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGTAC AAAAAGCCAG AAGACCTCTA CACCCGCGCT
1651 GGACAAGGCG AATTTCCAA CCGGCCATC CAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCGTAC CCGTCAGCGA AACCACCATC GTCAACAGT
1751 CCAAAATCAA AAAAGGCGGC AAAACGGCG TGCTCATCGA CGCGCAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGCTGC AAACCGCGC CGCCGACGA
1851 TATTATCGGC TTGTTACCC GCGAGCGCGG CATTTAGTG CACCGCAAAA
1901 CCTGCCCGTC TTTCCAACAC CTCGCCGAAC ACGCGCCCGA AAAAGTGCTG
1951 GACGCAAGCT GGGCGGCATT GCAGGAAGGA CAAGTATTCG CCGTCGATAT
2001 CGAAATCCGC GCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAAACTC AACGTTACCG CCGTGCAAAC CCAATCCCGC
2101 GACTTGGGAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTCAACGA
2151 CCTCCCGCGC GTCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201 TTACCCGCT TAA

```

This corresponds to the amino acid sequence <SEQ ID 420; ORF 117-1>:

```

m117-1.pep
1  MTAISPIQDT QSATLQELRE WFDSYCAALP DNDKNLIGTA WLLAQEHYPA
51  DAATPYGEPL PDHFLGAAQM VHELDLLPDA VAATLLADIG RYVPDNNLLV
101 SERCNSTVAE LVKGVDEVOK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRTLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQKPEKY REIALLLDEK RTERLEYIEN FLNILRGELK
251 KYNVHFEVAG RPKHIYSIYK KMKVKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDYI ANPKNGYKYS LHTVIVGPED KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401 LAAAFKTELF NDTIYVLTPH GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQORVE IITAKEGHPS VNWLYEGWVK SNKAIGKIRA
501 YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 QQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601 GLMTTLAKCC KPAPPDDIIG FVTRERGISV HRKTCPSFOH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRETL EVKQVNDLPR VLASLGDVKG VLSVTRL*

```

m117-1/g117-1 98.2% identity in 737 aa overlap

```

          10      20      30      40      50      60
m117-1.pep MTAISPIQDTQSATLQELREWFD SYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL
          |||
g117-1      MTAISPIQDTQSATLQELREWFD SYCAALPDNDKNLIGTAWLLAQEHYPADAATPYGEPL
          10      20      30      40      50      60
          70      80      90     100     110     120

```

sp|P55133|RELA_VIBSS GTP PYROPHOSPHOKINASE (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE) (PPGPP SYNTHETASE I) >gi|537617 (U13769) ppGpp synthetase I (Vibrio sp.) Length = 744

Score = 536 bits (1366), Expect = e-151

Identities = 288/685 (42%), Positives = 432/685 (63%), Gaps = 31/685 (4%)

Query: 74 LDLLPDAVAATLLADI---GRYVPDWNLLVSERCNSTVAELVKGVDVQKLTHFARVDSL 130
L + D + A LL + G Y D + E + T+ LV+GV+++ ++ ++ S
Sbjct: 68 LSMDDADTLIAALLYPLVEGGCYSTD---ALKEEYSGTILHLVQGVQMC AIS---QLKST 121

Query: 131 ATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRTLQFLSNAPDSPEKRAVAKETLDI 190
A +A Q + +R+MLL+MV D R V+IKLA R L+ + + PD +RA A+E +I
Sbjct: 122 AEETAQAAQVDNIRRMLLSMVDDFRCVVIKLAERICNLREVKDQPDDEV-RAAAQECANI 180

Query: 191 FAPLANRLGVWQLKWQLEDLGRHQKPEKYREIALLLDEKRTLEYIENFLNIRLGELK 250
+APLANRLG+ QLKW++ED FR+Q P+ Y++IA L E+R +R +YI +F++ L +K
Sbjct: 181 YAPLANRLGIGQLKWEIEDYAFRYQHPDPTKYIAKQLSERRIDREDYITHFVDDLSDAMK 240

Query: 251 KYNVHFEVAGRPKHIYSIYKMKVKKLSFDGLFDIRAVRILVDTVPECYTTLGIVHSLWQ 310
N+ EV GRPKHIYSI++KM KK L FD LFD+RAVRI+ + + +CY LG+VH+ ++
Sbjct: 241 ASNIRAQVQGRPKHIYSIWRKMQKKSLEFDELFDVRAVRIVAELQDCYAALGVVHTKYR 300

Query: 311 PIPGEFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTMDHQFNEFGVAHWRKYKEG- 369
+P EFDDY+ANPK NGY+S+HTV++GPE K +E+QIRT MH+ +E GVAHWR+YKEG
Sbjct: 301 HLPKEFDDYVANPKNGYQSIHTVVVGPEGKTIEIQTIRTKMHSESELGVAHWRKYKEGT 360

Query: 370 --GKGDSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPhgkVLSLP 427
G SAY++KI WLR+LL W+E M++SG ++ +++++F+D +Y TP G V+ LP
Sbjct: 361 ASGGQASAYDEKINWLRKLLAWQEEMSDSG--EMLDELRSQVFDREVYAFTPKGDDVVDLP 418

Query: 428 TGATPIDFAYALHSSIGDRCGAKVEGQIVPLSTPLENGQORVEIITAKEGHPSVNWLYE- 486
+ATP+DFAY +HS +G RC GAKVEG+IVP + L+ G +VEIIT KE +PS +WL
Sbjct: 419 SNATPLDFAYHIHSEVGHRCIGAKVEGRIVPFTYHLQMGDQVEIITQKEPNPSRDWLNPN 478

Query: 487 -GWVKSNAIGKIRAYIRQONADTVREGRVOLDKQLAKL--TPKPNLQELAEENLGYYKKP 543
G+V S++A K+ A+ R+Q+ D G+ L+ +L K+ T K A+ K P
Sbjct: 479 LGFVTSSRAKRVHAWFRKQDRDKNIIAGKEILEAEVLKIHATLKDAQYYAAKRFNVKSP 538

Query: 544 EDLYTAVGQGEIS-NRAIQACGTLNEPPVPVSETTIVKQSKI-----KKGKNGV 594
E+LY +G G++ N+ I +N+P + + K S+ KK ++ V
Sbjct: 539 EELYAGIGSGDLRINQVINHINALVNKPTAEEDQQLLEKLSEASNKQATSHKKPQRDAV 598

Query: 595 LIDGEDGLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPKVLDAW 654
+++G D LMT LA+CC+P P DDI GFVT+ RGISVHR C + L HAPE+++D W
Sbjct: 599 VVEGVDNLMTHLARCCQPIPGDDIQGFVTQGRGISVHRMDCEQLEELRHHAPERIIDTVW 658

Query: 655 AALQEGQVFAVDIEIRAQDRSGLLRDVSDALARHKLNVTAQTQ--SRDLEASMRFTLEV 712
G + + + + A +R+GLL++++ L K+ V ++++ + + M F LE+
Sbjct: 659 GGGFVGN-YTITVVRTASERNGLLKELTNTLMNEKVKVAGMKSRVDYKKQMSIMDFEEL 717

Query: 713 KQVNDLPRVLASLGDVKGVLVTRL 737
+ L RVL + VK V RL
Sbjct: 718 TDLEVLGRVLKRIEQVKDVAEKRL 742

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 421>:

a117-1.seq

```

1  ATGACCGCCA TCAGCCCGAT TCAAGACACG CAAAGCGCGA CTCTGCAAGA
51  ATTGCGCGAA TGGTTCGACA GCTACTGCAC CGCGCTGCCG AACAAACGATA
101  AAAAACTTGT CTTAGCCGCC CGTTCGCTGG CGGAAGCACA TTACCCCGCC
151  GATGCCGCCA CGCCGATGCG CGAACCGCTG CCCGACCACT TCCTCGGCGC
201  GCGCGAAATG GTTCATGAAC TCGACCTGCT CCCGATGCC GTCGCCGCCA
251  CCTGCTTGC CGACATCGGA CGCTACGTCC CCGACTGGAA CCTATTGGTT
301  TCCGAACGCT GCAACAGTAC CGTCGCCGAG CTGGTCAAAG GTGTGGACGA
351  AGTGACAGAA CTCACCCACT TCGCCCGGGT GGACAGCCTC GCCACGCGCG
401  AAGAACGCGC CCAGCAGGCA GAAACTATGC GGAAAATGCT GCTGGCGATG
451  GTTACCGACA TCCGCGTCGT GTTAATCAAA CTGGCGATGC GTACGCGCAC
501  CCTGCAATTT TTAAGCAACG CCCCAGACAG CCCCAGAAAA CGCGCCGTCG
551  CCAAAGAAAC CCTCGACATC TTCGCCCCGC TCGCCAACCG TTTGGGCGTG
601  TGGCAGCTCA AATGGCAGCT CGAAGATTG GGCTCCGCC ATCAAGAAC
651  CGAAAAATAC CGCGAAATCG CCCTGCTTTT GGACGAAAAA CGCACCGAAC
701  GCCTCGAATA CATCGAAAC TTCCTTAATA TCCTGCGTAC GGAACCAAA
751  AAATACAATA TCCACTTTGA AGTCGCCGGC CGTCCGAAAC ACATCTACTC
801  CATTACAAA AAAATGGTGA AGAAAAAAGT CAGCTTCGAC GGGTTGTTTCG

```

m117-1.pep PDHFLGAAQMVHEDLLPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
|||||
g117-1 PDHFLGAAQMVDELDPDAVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK
70 80 90 100 110 120

m117-1.pep 130 140 150 160 170 180
LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK
|||||
g117-1 LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKAMRTRTLQFLSNAPDSPEK
130 140 150 160 170 180

m117-1.pep 190 200 210 220 230 240
RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQKPEKYREIALLLDEKRTERLEYIEN
|||||
g117-1 RAVAKETLDIFAPLANRLGVWQLKWQLEDLGFRHQEPEKYREIALLLDEKRTERLEYIEN
190 200 210 220 230 240

m117-1.pep 250 260 270 280 290 300
FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKLSFDGLFDIRAVRILVDTVPECYT
||:|||||
g117-1 FLDILATELKKYNIHFEVAGRPKHIYSIYKKMVKKLSFDGLFDIRAVRILVDTVPECYT
250 260 270 280 290 300

m117-1.pep 310 320 330 340 350 360
TLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV
|||||
g117-1 TLGIVHSLWQPIPGFDDYIANPKGNGYKSLHTVIVGPEEKGVEVQIRTFDMHQFNEFGV
310 320 330 340 350 360

m117-1.pep 370 380 390 400 410 420
AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH
|||||
g117-1 AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH
370 380 390 400 410 420

m117-1.pep 430 440 450 460 470 480
GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS
|||||
g117-1 GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS
430 440 450 460 470 480

m117-1.pep 490 500 510 520 530 540
VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
|||||
g117-1 VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY
490 500 510 520 530 540

m117-1.pep 550 560 570 580 590 600
KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED
|||||
g117-1 KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSATTIVKQSKIKKGGKTGVLIDGED
550 560 570 580 590 600

m117-1.pep 610 620 630 640 650 660
GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDAWALQEG
|||||
g117-1 GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFRHLAEHAPEKVLDAWALQEG
610 620 630 640 650 660

m117-1.pep 670 680 690 700 710 720
QVFAVDIEIRAQDRSGLLRDVS DALARHKLNVAVQTQSRDLEASMRTLEVQVNDLPR
|||||
g117-1 QVFAVDIEIRAQDRSGLLRDVS DALARHKLNVAVQTQSRDLEASMRETLEVQVNDLPR
670 680 690 700 710 720

m117-1.pep 730
VLASLGDVKGVL SVTRLX
||:|||||
g117-1 VLAGLGDVKGVL SVTRLX
730

m117-1/RelA

```

851 ACATCCGCGC CGTGGCGGATT CTGGTTGATA CCGTCCCCGA GTGTTACACC
901 ACACTGGGCA TTGTCCACAG CCTCTGGCAG CCCATTCCCG GCGAGTTCGA
951 CGACTACATC GCCAACCCGA AAGGCAACGG CTATAAAAGT TTGCACACCG
1001 TCATCGTCGG CCCGGAAGAC AAAGGCGTGG AAGTGCAAAT CCGCACCTTC
1051 GATATGCACC AATTCAACGA ATTCGGTGTC GCCGCGCACT GCGGTATCAA
1101 AGAGGGCGGC AAAGGCGATT CCGCCTACGA ACAAAAAATC GCCTGGTTAC
1151 GCCAACTTTT GGAATGGGCG GAAACATGG CCGAAAGCGG CAAGGAAGAC
1201 CTCGCCGCGC CCTTCAAACG CGAGCTTTTC AACGACACGA TTTATGTTTT
1251 GACCCCGCAC GGCAGAGTCC TCTCCCTGCC CACAGGCGCG ACCCCCATCG
1301 ACTTCGCCTA CGCCCTGCAC AGCAGCATCG GCGACCGTTG CCGCGGTGCG
1351 AAAGTCGAAG GGCAGATTGT GCCGCTGTCC ACCCCGCTCG AAAACGGACA
1401 GCGGTGTCGAA ATCATTACCG CCAAAGAAGG GCATCCTTCC GTCAACTGGC
1451 TTTACGAAGG CTGGGTCAAA TCCAACAAGG CAATCGGCAA AATCCGCGCC
1501 TACATCCGCC AGCAAAACGC CGACACCGTG CGCGAAGAAG GCCGCGTCCA
1551 ACTCGACAAA CAGCTTGCCA AACTCACGCC CAAACCCAAC CTGCAAGAGC
1601 TTGCCGAAAA TCTCGGCTAC AAAAAGCCAG AAGACCTCTA CACCGCCGTC
1651 GGACAAGGCG AAATTTCCAA CCGCGCCATC CAAAAAGCCT GCGGCACGCT
1701 GAACGAACCG CCGCCCGTAC CCGTCAGCGA AACCACCATC GTCAACACGT
1751 CAAAAATCAA AAAAGGCGGC AAAAACGGCG TGCTCATCGA CGGCGAAGAC
1801 GGTCTGATGA CCACGCTTGC CAAATGTGTC AAACCCGCGC CGCCCGACGA
1851 CATGTGCGGC TTCGTTACCC GCGATCGCGG CATTTCGGTA CACCGCAAAA
1901 CCTGCCCCCTC TTTCCGACAC CTCGCCGAAC ACGCGCCCGA AAAAGTACTG
1951 GACGCAAGTT GGGCGGCGTT GCAGGAAGGA CAAGTGTTCG CCGTCGATAT
2001 CGAAATCCGC GCCCAAGACC GCTCCGGGCT TTTGCGCGAC GTATCCGACG
2051 CGCTCGCCCG CCACAACTC AACGTTACCG CCGTGCAAAAC CCAGTCCCGC
2101 GACTTGGAAG CCAGCATGAG GTTCACGCTC GAAGTCAAAC AAGTTACCGA
2151 CCTCCACGCG GTCCTCGCCA GCCTCGGCGA CGTCAAAGGC GTATTGAGCG
2201 TTACCCGCGC TTAA

```

This corresponds to the amino acid sequence <SEQ ID 422; ORF 117-1.a>:

a117-1.pep

```

1  MTAISPIQDT QSATLQELRE WFDSYCTALP NNDKKLVLA RSLAEAHYPA
51  DAATPYGEPL PDHFLGAQM VHELDLLPDA VAATLLADIG RYVPDWNLLV
101 SERCNSTVAE LVKGVDEVQK LTHFARVDSL ATPEERAQQA ETMRKMLLAM
151 VTDIRVVLIK LAMRTRLQF LSNAPDSPEK RAVAKETLDI FAPLANRLGV
201 WQLKWQLEDL GFRHQEPEKY REIALLLDEK RTERLEYIEN FLNILRTELK
251 KYNHFEVAG RPKHIYSIYK KMKVKKLSFD GLFDIRAVRI LVDTVPECYT
301 TLGIVHSLWQ PIPGEFDDYI ANPKNGYKKS LHTVIVGPED KGVEVQIRTF
351 DMHQFNEFGV AAHWRYKEGG KGDSAYEQKI AWLRQLLDWR ENMAESGKED
401 LAFAFKTELF NDTIYVLTPL GKVLSLPTGA TPIDFAYALH SSIGDRCRGA
451 KVEGQIVPLS TPLENGQVRE IITAKEGHPV VNWLYEGWVK SNKAIGKIRA
501 YIRQONADTV REEGRVQLDK QLAKLTPKPN LQELAENLGY KKPEDLYTAV
551 GQGEISNRAI QKACGTLNEP PPVPVSETTI VKQSKIKKGG KNGVLIDGED
601 GLMTTLAKCC KPAPPDDIVG FVTRDRGISV HRKTCPSFRH LAEHAPEKVL
651 DASWAALQEG QVFAVDIEIR AQDRSGLLRD VSDALARHKL NVTAVQTQSR
701 DLEASMRFTL EVKQVTDLPR VLASLGDVKG VLSVTRL*

```

a117-1/m117-1 97.7% identity in 737 aa overlap

	10	20	30	40	50	60
m117-1.pep	MTAISPIQDTQSATLQELREWFDSCAALPDNDKNLIGTAWLLAQEHYPADATPYGEPL					
a117-1	MTAISPIQDTQSATLQELREWFDSCAALPDNDKNLIGTAWLLAQEHYPADATPYGEPL					
	10	20	30	40	50	60
m117-1.pep	PDHFLGAQMVHELDLLPDVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK					
a117-1	PDHFLGAQMVHELDLLPDVAATLLADIGRYVPDWNLLVSERCNSTVAELVKGVDEVQK					
	70	80	90	100	110	120
m117-1.pep	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRLQFLSNAPDSPEK					
a117-1	LTHFARVDSLATPEERAQQAETMRKMLLAMVTDIRVVLIKLAMRTRLQFLSNAPDSPEK					
	130	140	150	160	170	180
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGRHQEPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGRHQEPEKYREIALLLDEKRTERLEYIEN					
	190	200	210	220	230	240
m117-1.pep	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGRHQEPEKYREIALLLDEKRTERLEYIEN					
a117-1	RAVAKETLDIFAPLANRLGVWQLKWQLEDLGRHQEPEKYREIALLLDEKRTERLEYIEN					
	190	200	210	220	230	240

339

m117-1.pep	250	260	270	280	290	300
	FLNILRGELKKYNVHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT					
a117-1	250	260	270	280	290	300
	FLNILRTELKKYNHFEVAGRPKHIYSIYKKMVKKKLSFDGLFDIRAVRILVDTVPECYT					
m117-1.pep	310	320	330	340	350	360
	TLGIVHSLWQPIPGFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV					
a117-1	310	320	330	340	350	360
	TLGIVHSLWQPIPGFDDYIANPKNGYKSLHTVIVGPEDKGVEVQIRTFDMHQFNEFGV					
m117-1.pep	370	380	390	400	410	420
	AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH					
a117-1	370	380	390	400	410	420
	AAHWRYKEGGKGSAYEQKIAWLRQLLDWRENMAESGKEDLAAAFKTELFNDTIYVLTPH					
m117-1.pep	430	440	450	460	470	480
	GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS					
a117-1	430	440	450	460	470	480
	GKVLSLPTGATPIDFAYALHSSIGDRCRGAKVEGQIVPLSTPLENGQORVEIITAKEGHPS					
m117-1.pep	490	500	510	520	530	540
	VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY					
a117-1	490	500	510	520	530	540
	VNWLYEGWVKSNAIGKIRAYIRQQNADTVREEGRVQLDKQLAKLTPKPNLQELAENLGY					
m117-1.pep	550	560	570	580	590	600
	KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED					
a117-1	550	560	570	580	590	600
	KKPEDLYTAVGQGEISNRAIQKACGTLNEPPVPVSETTIVKQSKIKKGGKNGVLIDGED					
m117-1.pep	610	620	630	640	650	660
	GLMTTLAKCKPAPPDDIIGFVTRERGISVHRKTCPSFQHLAEHAPEKVLDASWAALQEG					
a117-1	610	620	630	640	650	660
	GLMTTLAKCKPAPPDDIVGFVTRDRGISVHRKTCPSFRHLAEHAPEKVLDASWAALQEG					
m117-1.pep	670	680	690	700	710	720
	QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVAVQTQSRDLEASMRETLEVQVNDLPR					
a117-1	670	680	690	700	710	720
	QVFAVDIEIRAQDRSGLLRDVSDALARHKLNVAVQTQSRDLEASMRETLEVQVNDLPR					
m117-1.pep	730					
	VLASLGDVKGVLVSVTRLX					
a117-1	730					
	VLASLGDVKGVLVSVTRLX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 423>:

```

g118.seq
1  ATGTGCGAGT TCAAGGATTT TAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGCTG GAGAATGATT TAATcgaGGT TAGGAGAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCAAATTG GGAGCTTTT GAAATTAAG
251 CTTCCCTTG GTTGCCTGAT AGCGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAaCA AAAAATAG

```

This corresponds to the amino acid sequence <SEQ ID 424; ORF 118.ng>:

```

g118.pep
1  MCFEKDFRRN IPCFEEYDEN SFIGKWYDDG VWDDEEYWL ENDLIEVRRK
51  YPYPM DIPRD IVIGIGTIID FLMVFNWELF EIKASPLWLPD SVGIHERYER
101 FTTMLRYIFT EKDIVNVRED YYNKK*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 425>:

```
m118.seq
1  ATGTGTGAGT TCAAGGATAT TATAAGAAAC GTTCCTTATT TTGAGGGGTA
51  TGACGAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAAGTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
151 TATCCTTATC CGATGGACAT ACCAAGATAT GTTGTGATTG GAATCGGTAC
201 CATATTGAT TTCTTAATGG TTCCAAATG GAAACTTTT GAAATTAAAG
251 CTTCCCCTTG GTTGCCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAACA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 426; ORF 118>:

```
m118.pep
1  MCEF KDII RN VPF EGYDEN SFI GKWYDDG VWD DEEYWK L ENDLIEVRKK
51  YPYP MDIPRY VVIGIGTIID FLMV PNWKLF EIKAS PWLPD SVGIHERYER
101 FTTMLRYIFT EKDIVNVRF D YYNKK*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 118 shows 92.8% identity over a 125 aa overlap with a predicted ORF (ORF 118.ng) from *N. gonorrhoeae*:

```
m118/g118

          10          20          30          40          50          60
m118.pep  MCEF KDII RNVPYF EGYDENSFI GKWYDDG VWD DEEYWK L ENDLIEVRKKYPYPMDIPRY
          |||||: ||:| || |||||
g118       MCEF KD FRRNIPCF EGYDENSFI GKWYDDG VWD DEEYWK L ENDLIEVRKKYPYPMDIPRD
          10          20          30          40          50          60

          70          80          90          100         110         120
m118.pep  VVIGIGTIIDFLMVPNWKLF EIKAS PWLPD SVGIHERYERFTTMLRYIFTEKDIVNVRF D
          :|||||
g118       IVIGIGTIIDFLMVPNWKLF EIKAS PWLPD SVGIHERYERFTTMLRYIFTEKDIVNVRF D
          70          80          90          100         110         120

m118.pep  YYNKKX
          |||||
g118       YYNKKX
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 427>:

```
a118.seq
1  ATGTGTGAGT TCAAGGATT TTAGAAGAAAC ATCCCTTGTT TTGAAGAGTA
51  TGACGAAAT TCATTTATTG GCAAATGGTA TGATGACGGG GTGTGGGATG
101 ATGAAGAATA TTGGAATTTG GAGAATGATT TAATCGAGGT TAGAAAAAAA
151 TATCCTTATC CGATGGATAT ACCAAGGGAT ATTGTGATTG GAATCGGTAC
201 CATTATTGAT TTTTAAATGG TTCCAAATTG GGAGCTTTT GAAATTAAAG
251 CTTCCCCTTG GTTGCCTGAT AGTGTGGGAA TTCATGAACG TTATGAAAGA
301 TTCACAACGA TGCTCCGTTA TATTTTACC GAGAAAGACA TAGTCAACGT
351 GCGATTTGAT TATTACAACA AAAAATAG
```

This corresponds to the amino acid sequence <SEQ ID 428; ORF 118.a>:

```
a118.pep
1  MCEF KD FRRN IPCF EGYDEN SFI GKWYDDG VWD DEEYWK L ENDLIEVRKK
51  YPYP MDIPRD IVIGIGTIID FLMV PNWELF EIKAS PWLPD SVGIHERYER
101 FTTMLRYIFT EKDIVNVRF D YYNKK*
```

m118/a118 93.6% identity in 125 aa overlap

```
m118.pep  MCEF KDII RNVPYF EGYDENSFI GKWYDDG VWD DEEYWK L ENDLIEVRKKYPYPMDIPRY
          |||||: ||:| || |||||
a118       MCEF KD FRRNIPCF EGYDENSFI GKWYDDG VWD DEEYWK L ENDLIEVRKKYPYPMDIPRD
          10          20          30          40          50          60
```

341

	70	80	90	100	110	120
m118.pep	VVIGIGTIIDFLMVPNWKLF EIKAS PWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRED					
a118	IVIGIGTIIDFLMVPNWELFEIKAS PWLPDSVGIHERYERFTTMLRYIFTEKDIVNVRED					
	70	80	90	100	110	120
m118.pep	YYNKKX					
a118	YYNKKX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 429>:

```

g120.seq
1  ATGATGAAGA CTTTAAAAA TATATTTTCC GCCGCCATT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAAGGCTACC CCAATCCGCC GTGCTGCACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201 TTTCGAATCC GCGGGTACGG TTGTCGGCAA TACCCTGCAC CCTGCCTACT
251 ATAAAGACAT ACGCAGGGGC AAATCGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC
351 CAAGGCTATG GATTTGTTCA CGCTTGCTTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGCC TGAATAAGGC GGGTACGGGA AAATACAGCA Taggcggcgt
501 gGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATACGGTAA
551 CGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGAcgaCG GCAAAACCTA TACGCTGAAG CTCAAATCGG TGCAGATCAA
651 CGGACAGGCC GCCAAACCGT AA

```

This corresponds to the amino acid sequence <SEQ ID 430; ORF 120.ng>:

```

g120.pep
1  MMKTFKNIFS AAILSALPC AYAARLPQSA VLHYSGSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNLTH PAYYKDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPLG KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DTVTYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 431>:

```

m120.seq
1  ATGATGAAGA CTTTAAAAA TATATTTTCC GCCGCCATT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGmACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201 TTTCGAGTCC GCGGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
251 ATAGAGACAT ACGCAGGGGC AAATCGTATG CGGAAGCcAA ATTCGCCGAC
301 GGCAGCGTAA CTTACGGCAA AGCGGGCGAG AGCAAAACCG AGCAAAGCCC
351 CAAGGCTATG GATTTGTTCA CGCTTGCTTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGTT TGAATAAGGC GGGTACAGGA AAATACAGCA TAGGCGGCGT
501 GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
551 TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGACGACG GCAAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
651 CGGCCAGGCA GCCAAACCG

```

This corresponds to the amino acid sequence <SEQ ID 432; ORF 120>:

```

m120.pep
1  MMKTFKNIFS AAILSALPC AYAAGLPQSA VLXYSYSYGI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNLTH PTYYRDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPLG KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAQIGY
201 TDDGKTYTLK LKSVQINGQA AKP

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 120 shows 97.3% identity over a 223 aa overlap with a predicted ORF (ORF 120.ng) from *N. gonorrhoeae*:

m120/g120

	10	20	30	40	50	60
m120.pep	MMKTFKNIFSAAILSAALPCAYAAGLPQSAVLXYSYGI	PATMTFERSGNAYKIVSTIK				
g120	MMKTFKNIFSAAILSAALPCAYAARLPQSAVLHYSYGI	PATMTFERSGNAYKIVSTIK				
	10	20	30	40	50	60
	70	80	90	100	110	120
m120.pep	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGS	VTYKGAGESKTEQSPKAM				
g120	VPLYNIRFESGGTVVGNTLHPAYYKDIRRGKLYAEAKFADGS	VTYKGAGESKTEQSPKAM				
	70	80	90	100	110	120
	130	140	150	160	170	180
m120.pep	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETE	VVKYRVRRGD				
g120	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETE	VVKYRVRRGD				
	130	140	150	160	170	180
	190	200	210	220		
m120.pep	DAVMYFFAPSLNNIPAIQIGYTDDGKTYTLKLSVQINGQA	AKP				
g120	DTVTYFFAPSLNNIPAIQIGYTDDGKTYTLKLSVQINGQA	AKPX				
	190	200	210	220		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 433>:

a120.seq

```

1  ATGATGAAGA CTTTAAAAA TATATTTTCC GCCGCCATTT TGTCCGCCGC
51  CCTGCCGTGC GCGTATGCGG CAGGGCTGCC CCAATCCGCC GTGCTGCACT
101 ATTCCGGCAG CTACGGCATT CCCGCCACGA TGACATTTGA ACGCAGCGGC
151 AATGCTTACA AAATCGTTTC GACGATTAAA GTGCCGCTAT ACAATATCCG
201 TTTCGAGTCC GCGGTACGG TTGTCGGCAA TACCCTGCAC CCTACCTACT
251 ATAGAGACAT ACGCAGGGGC AAATGTATG CGGAAGCCAA ATTCGCCGAC
301 GGCAGCGTAA CCTACGGCAA AGCGGGCGAG AGCAAACCC AGCAAAGCCC
351 CAAGGCTATG GATTGTGTTCA CGCTTGCCCTG GCAGTTGGCG GCAAATGACG
401 CGAAACTCCC CCCGGGGCTG AAAATCACCA ACGGCAAAAA ACTTTATTCC
451 GTCGGCGGTT TGAATAAGGC GGGTACAGGA AAATACAGCA TAGGCGGCGT
501 GGAAACCGAA GTCGTCAAAT ATCGGGTGCG GCGCGGCGAC GATGCGGTAA
551 TGTATTTCTT CGCACCGTCC CTGAACAATA TTCCGGCACA AATCGGCTAT
601 ACCGACGACG GCAAACCTA TACGCTGAAA CTCAAATCGG TGCAGATCAA
651 CGGCCAGGCA GCCAAACCGT AA

```

This corresponds to the amino acid sequence <SEQ ID 434; ORF 120.a>:

a120.pep

```

1  MMKTFKNIFS AAILSAALPC AYAAGLPQSA VLHYSYGYI PATMTFERSG
51  NAYKIVSTIK VPLYNIRFES GGTVVGNTLH PTYYRDIRRG KLYAEAKFAD
101 GSVTYGKAGE SKTEQSPKAM DLFTLAWQLA ANDAKLPPGL KITNGKKLYS
151 VGGLNKAGTG KYSIGGVETE VVKYRVRRGD DAVMYFFAPS LNNIPAIQIGY
201 TDDGKTYTLK LKSVQINGQA AKP*

```

m120/a120 99.6% identity in 223 aa overlap

	10	20	30	40	50	60
m120.pep	MMKTFKNIFSAAILSAALPCAYAAGLPQSAVLXYSYGI	PATMTFERSGNAYKIVSTIK				
a120	MMKTFKNIFSAAILSAALPCAYAAGLPQSAVLHYSYGI	PATMTFERSGNAYKIVSTIK				

343

	10	20	30	40	50	60
m120.pep	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVYTGKAGESKTEQSPKAM					
a120	70	80	90	100	110	120
	VPLYNIRFESGGTVVGNTLHPTYYRDIRRGKLYAEAKFADGSVYTGKAGESKTEQSPKAM					
m120.pep	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEYVVKYRVRGD					
a120	130	140	150	160	170	180
	DLFTLAWQLAANDAKLPPGLKITNGKKLYSVGGLNKAGTGKYSIGGVETEYVVKYRVRGD					
m120.pep	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX					
a120	190	200	210	220		
	DAVMYFFAPSLNNIPAQIGYTDDGKTYTLKLKSVQINGQAAKPX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 435>:

g121.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATTATG TCGGGAACCA GTATGGACGG
51  GCGGATGCC GTGCTGGTAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCTGACC GGTTCGCGCG CAAATTGCTG
151 GATTTCAGG ACACAGGCAC AGACGAATG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCTCCGTGC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTtac AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACtGa cgcggatttT TACCGTCggc gacttcCGCA
401 GCCGCGACCT TGCTGCCGCG GGacaAGGTG CGCCGCTCGT CCCCgcCTTT
451 CACGAAGCCC TGTTCGCGA TGACAGGGAA ACACGCGTGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCggCGCA CCGCCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGAcgcgtg gacgcaggca
601 cacTGGcagc TGCCTTACGA CAAAAacggt gcAAAGgcgg cacAAGGCAA
651 catatTGCcg cAACTGCTCG gcaggctGCT CGCCcaccCG TATTTCTCAC
701 AACCcacc ccAAAGCACG GGgcGCGaac Tgtttgccct AAattggctc
751 gaaacctAcc ttgacggcgg cgaaccgga tacgacgtat tgcggacgct
801 ttcccgattc accgcgcaaA ccgTttggga cgccgtctca CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCTCAATG GGTGGAGGCG gccgCATTtg
1001 cgtggttgC GCGGTGTTGG ATTAACCGCA TTCCCGTAG TCCGCACAAA
1051 GCGACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 436; ORF 121.ng>:

g121.pep

```

1  METQLYIGIM SGTSMDGADA VLVRMDGGKW LGAEGHAFTF YPDRLRRKLL
51  DLQDTGTDEL HRSRMLSQEL SRLYAQTAAE LLCSONLAPC DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL AELTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRVVLNIGGI ANISVLPPGA PAFGFDTPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLLGRLLAHP YFSQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVWDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWLAACW INRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 437>:

m121.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCGC CCAATTGCTG

```

344

```

151 GATTTCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTGTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCACCCTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351 GCCGCTGCTG GCGxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
401 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
451 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
501 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
551 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
601 xxxxxxGAGC TTCCTTACGA CAAAACGGT GCAAAGTCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTCGCAC
701 AACGCCACCC TAAAGCACG GGGCGCGAAC TGTTTGCCAT AAATTGGCTC
751 GAAACCTACC TTGACGCGCG GAAAACCGA TACGACGTAT TGCGGACGCT
801 TTCCCGTTTT ACCGCGCAA CCGTTTGCGA CGCCGTCTCA CACGCGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGACG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCGGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGnATTG
1001 CGTGGTTGGC GCGGTGTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGGCG CATCAAACC GTGTATTCTG AnCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 438; ORF 121>:

m121.pep

```

1 METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRQLL
51 DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPEHGY SIQLADLPLL Axxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
151 xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx xxxxxxxxxxx
201 xxQLPYDKNG AKSAQGNILP QLLDRLLAHP YFAQRHPKST GRELFAINWL
251 ETYLDGGENR YDVLRLTSRF TAQTVCDAYS HAAADARQMY ICDGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPQWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

Computer analysis of the amino acid sequences gave the following results:

Homology with a predicted ORF from *N. meningitidis* *menA* with *menB*

ORF 121 shows 73.5% identity over a 366 aa overlap with a predicted ORF (ORF121.ng) from *N. gonorrhoeae*:

m121/g121

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTYPGRLLRQLLDLQDTGADEL					
g121	METQLYIGIMSGTSMGDADAVLVRMDGGKWLGAEGHAFTYPDRLLRRLKLLDLQDTGTDEL					
	10	20	30	40	50	60
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAAELLCSONLAPSDITALGCHGQTVRHAPEHGYISQLADLPLL					
g121	HRSRMLSQELSRLYAQTAAELLCSONLAPCDITALGCHGQTVRHAPEHGYISQLADLPLL					
	70	80	90	100	110	120
	130	140	150	160	170	180
m121.pep	AXXX					
g121	AELTRIFTVGDFRSRDLAAGGQGAPLVPFHEALFRDDRETRVVLNIGGIANISVLPPGA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
g121	PAFGFDTGFGNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLGRLLAHPYFSQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYLDGGENRYDVLRLTSRFTAQTVCDAYSHAAADARQMYICDGGIRNPV					
g121	GRELFALNWLETYLDGGENRYDVLRLTSRFTAQTVWDAYSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300

345

	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTADLNLDPOQWVEAAFAWLAACWINRIPGSPHKATGASKPCIL					
g121	LMADLAECFGTRVSLHSTAE LNLDPOQWVEAAFAWLAACWINRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYYX					
g121	GAGYYYYX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 439>:

```

a121.seq
1   ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GCGCGATGCC GTACTGATAC GGATGGACGG CGGCAAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTACGCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAATG CACCGCAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCCGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGCTC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TGCGGCCGCG GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCGCGCA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCACGCA CCGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCG TATTTCCGAC
701 AACCACCACC TAAAGCACG GGGCGCGAAC TGTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGGCGG CGAAAACCGA TACGACGTAT TCGGGACGCT
801 TTCCCGATT ACCGCGCAA CCGTTTCGA CGCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACGCGCGAA CTGAACCTCG ATCCGCAATG GGTAAGGCC GCGCGTTCG
1001 CATGGATGGC GCGGTGTTGG GTCAACCGCA TTCCCGTAG TCCGCACAAA
1051 GCAACCGGCG CATCCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 440; ORF 121.a>:

```

a121.pep
1   METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRLLRRKLL
51  DLQDTGADEL HRSRMLSQEL SRLYAQTAAE LLCSQNLAPS DITALGCHGQ
101 TVRHAPESY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWMQA
201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAE LNLDPOQWVEA AAFAWMAACW VNRIPGSPHK
351 ATGASKPCIL GAGYYY*

```

m121/a121 74.0% identity in 366 aa overlap

	10	20	30	40	50	60
m121.pep	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
a121	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
	10	20	30	40	50	60
m121.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEHYSYIQLADLPLL					
a121	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEHYSYIQLADLPLL					
	70	80	90	100	110	120
m121.pep	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEHYSYIQLADLPLL					
a121	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPHEHYSYIQLADLPLL					
	70	80	90	100	110	120

346

	130	140	150	160	170	180
m121.pep	XX					
a121	AERTQIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
	190	200	210	220	230	240
m121.pep	XXXXXXXXXXXXXXXXXXXXXXXXXQLPYDKNGAKSAQGNILPQLLDRLLAHPYFAQRHPKST					
a121	PAFGFDTGPGNMLMDAWMOAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
	250	260	270	280	290	300
m121.pep	GRELFAINWLETYLDGGENRYDVLRLTSRFTAQTVCDVASHAAADARQMYICDGGIRNPV					
a121	GRELFAINWLETYLDGGENRYDVLRLTSRFTAQTVFCDVASHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
	310	320	330	340	350	360
m121.pep	LMADLAECFGTRVSLHSTADLNLDPWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAEINLDPWVEAAAFAWMAACWVNRIIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121.pep	XAGYYYYX					
a121	GAGYYYYX					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 441>:

m121-1.seq

```

1  ATGGAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGC GGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTTCAGCCG CCAATTGCTG
151 GATTTGCAGG ACACAGGCGC AGACGAACTG CACCGCAGCA GGATTTTGTC
201 GCAAGAACTC AGCCGCCTAT ATGCGCAAAAC CGCGCGCGAA CTGCTGTGCA
251 GTCAAACCTC CGCACCGTCC GACATTACCG CCCTCGGCTG CCACGGGGCAA
301 ACCGTCCGAC ACGCGCCGGA ACACGGTTAC AGCATACAGC TTGCCGATTT
351 GCCGTGCTG GCGGAACGGA CGCGGATTTT TACCGTCGGC GACTTCCGCA
401 GCCCGACCTC TGCGGCGCGC GGACAAGGCG CGCCACTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CAACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCGACGCA CCCGCCTTCG
551 GCTTCGACAC AGGGCCGGGC AATATGCTGA TGGACGCGTG GACGCAAGCA
601 CACTGGCAGC TTCCTTACGA CAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCACCCCG TATTTCGCAC
701 AACCCACCCC TAAAGACAGC GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGCGCG CGAAAACCGA TACGACGTAT TCGGAGCGCT
801 TTCCCGTTTT ACCGCGCAAA CCGTTTGCGA CGCCGTCTCA CACGACGCGG
851 CAGATGCCCG TCAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAC CTGAACCTCG ATCCGCAATG GGTGGAAGCC GCCGNATTG
1001 CGTGTTGGC GGC GTTGG ATTAATCGCA TTCCCGGTAG TCCGCACAAA
1051 GCAACCGCGC CATCAAACC GTGTATTCTG ANCGCGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 442; ORF 121-1>:

m121-1.pep

```

1  METQLYIGIM SGTSMGADA VLIRMDGGKW LGAEGHAFTP YPGLRRQLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAAE LLCSONLAPS DITALGCHGQ
101 TVRHAPEHG Y SIQLADPLL AERTRIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDNRE TRAVLNIGGI ANISVLPPDA PAFGFDTGPG NMLMDAWTQA
201 HWQLPYDKNG AKAAQGNILP QLDRLLAHP YFAQPHPKST GRELFALNWL
251 ETYLDGGENR YDVLRLTSRF TAQTVCDAVS HAAADARQMY ICGGGIRNPV
301 LMADLAECFG TRVSLHSTAD LNLDPWVEA AXFAWLAACW INRIPGSPHK
351 ATGASKPCIL XAGYYY*

```

m121-1/g121 95.6% identity in 366 aa overlap

10 20 30 40 50 60

347

```

m121-1.pep  METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTYPGRRLRRQLLDLQDTGADEL
g121         METQLYIGIMSGTSMGDADAVLVRMDGGKWLGAEGHAFTYPDRRLRRKLLDLQDTGTDEL
              10      20      30      40      50      60

              70      80      90      100     110     120
m121-1.pep  HRSRILSQELSRLYAQTAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL
g121         HRSRILSQELSRLYAQTAELLCSQNLAPCDITALGCHGQTVRHAPEHGYSIQLADLPLL
              70      80      90      100     110     120

              130     140     150     160     170     180
m121-1.pep  AERTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDNRETRAVLNIGGIANISVLPPDA
g121         AELTRIFTVGDFRSRDLAAGGQGAPLVPAFHEALFRDDRETRVVLNIGGIANISVLPPGA
              130     140     150     160     170     180

              190     200     210     220     230     240
m121-1.pep  PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST
g121         PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLRLLAHPYFSQPHPKST
              190     200     210     220     230     240

              250     260     270     280     290     300
m121-1.pep  GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDASHAAADARQMYICGGGIRNPV
g121         GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVWDASHAAADARQMYICGGGIRNPV
              250     260     270     280     290     300

              310     320     330     340     350     360
m121-1.pep  LMADLAECFGTRVSLHSTADLNDLPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL
g121         LMADLAECFGTRVSLHSTAE LNDLPQWVEAAAFWLAACWINRIPGSPHKATGASKPCIL
              310     320     330     340     350     360

m121-1.pep  XAGYYYY
g121         GAGYYYY

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 443>:

```

a121-1.seq
1  ATGGAAACAC AGCTTTACAT CGGCATCATG TCGGGAACCA GCATGGACGG
51  GGCGGATGCC GTACTGATAC GGATGGACGG CGGCAATGG CTGGGCGCGG
101 AAGGGCACGC CTTTACCCCC TACCCCGGCA GGTACGCCG CAAATTGCTG
151 GATTTCAGG ACACAGGCGC GGACGAACTG CACCCGAGCA GGATGTTGTC
201 GCAAGAACTC AGCCGCCTGT ACGCGCAAAC CGCCGCGGAA CTGCTGTGCA
251 GTCAAAACCT CGCGCCGTCC GACATTACCG CCCTCGGCTG CCACGGGCAA
301 ACCGTCAGAC ACGCGCCGGA ACACAGTTAC AGCGTACAGC TTGCCGATTT
351 GCCGCTGCTG GCGGAACGGA CTCAGATTTT TACCGTCGGC GACTTCCGCA
401 GCCGCGACCT TCGGCGCGGC GGACAAGGCG CGCCGCTCGT CCCCGCCTTT
451 CACGAAGCCC TGTTCCGCGA CGACAGGGAA ACACGCGCGG TACTGAACAT
501 CGGCGGGATT GCCAACATCA GCGTACTCCC CCCCAGCGCA CCGGCCTTCG
551 GCTTCGACAC AGGACCGGGC AATATGCTGA TGGACGCGTG GATGCAGGCA
601 CACTGGCAGC TTCCTTACGA CAAAAACGGT GCAAAGGCGG CACAAGGCAA
651 CATATTGCCG CAACTGCTCG ACAGGCTGCT CGCCCACCCG TATTTGCGAC
701 AACCCACCCC TAAAAGCACG GGGCGCGAAC TGTTTGCCCT AAATTGGCTC
751 GAAACCTACC TTGACGCGCG CGAAAACCGA TACGACGTAT TCGGGACGCT
801 TTCCCGATTG ACCGCGCAAA CCGTTTTCGA CGCCGTCTCA CACGCAGCGG
851 CAGATGCCCG TCAAATGTAC ATTTGCGGCG GCGGCATCCG CAATCCTGTT
901 TTAATGGCGG ATTTGGCAGA ATGTTTCGGC ACACGCGTTT CCCTGCACAG
951 CACCGCCGAA CTGAACCTCG ATCCGCAATG GGTAGAAGCC GCCGCGTTCG
1001 CATGGATGGC GGCGTGTGG GTCAACCGCA TTCCCGTAG TCCGCACAAA
1051 GCAACCGCGC CATCAAACC GTGTATTCTG GCGCGGGAT ATTATTATTG
1101 A

```

This corresponds to the amino acid sequence <SEQ ID 444; ORF 121-1.a>:

```

a121-1.pep
1  METQLYIGIM SGTSMGDADA VLIRMDGGKW LGAEGHAFTP YPGRRLRRKLL
51  DLQDTGADEL HRSRILSQEL SRLYAQTAEE LLCSQNLAPS DITALGCHGQ
101 TVRHAPEHSY SVQLADLPLL AERTQIFTVG DFRSRDLAAG GQGAPLVPAF
151 HEALFRDDRE TRAVLNIGGI ANISVLPPDA PAFGFDTPGPNMLMDAWMQA

```

348

201 HWQLPYDKNG AKAAQGNILP QLLDRLLAHP YFAQPHPKST GRELFALNWL
 251 ETYLDGGENR YDVLRTLSRF TAQTVFDAVS HAAADARQMY ICGGGIRNPV
 301 LMADLAECFG TRVSLHSTAE LNLDPQWVEA AAFAWMAACW VNRIPGSPHK
 351 ATGASKPCIL GAGYYY*

m121-1/a121-1 96.4% identity in 366 aa overlap

m121-1.pep	10	20	30	40	50	60
	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
a121-1	METQLYIGIMSGTSMGDADAVLIRMDGGKWLGAEGHAFTPYPGRLRRQLLDLQDTGADEL					
	10	20	30	40	50	60
m121-1.pep	70	80	90	100	110	120
	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHGYSIQLADLPLL					
a121-1	HRSRILSQELSRLYAQTAAELLCSQNLAPSDITALGCHGQTVRHAPEHSYSVQLADLPLL					
	70	80	90	100	110	120
m121-1.pep	130	140	150	160	170	180
	AERTRIFTVGDFFRSRDLAAGGQGAPLVPFHEALFRDNRETRAVLNIGGIANISVLPPDA					
a121-1	AERTQIFTVGDFFRSRDLAAGGQGAPLVPFHEALFRDDRETRAVLNIGGIANISVLPPDA					
	130	140	150	160	170	180
m121-1.pep	190	200	210	220	230	240
	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
a121-1	PAFGFDTPGPNMLMDAWTQAHWQLPYDKNGAKAAQGNILPQLLDRLLAHPYFAQPHPKST					
	190	200	210	220	230	240
m121-1.pep	250	260	270	280	290	300
	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVCDASHAAADARQMYICGGGIRNPV					
a121-1	GRELFALNWLETYLDGGENRYDVLRTLSRFTAQTVFDAVSHAAADARQMYICGGGIRNPV					
	250	260	270	280	290	300
m121-1.pep	310	320	330	340	350	360
	LMADLAECFGTRVSLHSTADLNLDPQWVEAAXFAWLAACWINRIPGSPHKATGASKPCIL					
a121	LMADLAECFGTRVSLHSTAE LNLDPQWVEA AAFAWMAACWVNRIPGSPHKATGASKPCIL					
	310	320	330	340	350	360
m121-1.pep	XAGYYYYX					
a121	GAGYYYYX					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 445>:

g122.seq

1	ATGGCTTTAC	TGAGCATCCG	CAAGCTGCAC	AAACAATACG	GCAGCGTAAC
51	CGCCATCCAA	TCCTTAGACT	TGGACTTGGG	AAAAGGCGAA	GtcatCGTAC
101	TGCTGGGCCC	gTccggetgc	ggCAAATCCA	CCCTcctgcg	ctgcgtaac
151	GGTTTGGAGC	CGCACCaaag	cgGCAGCATC	GTGATGGACG	GTgtcgGCGA
201	ATTGggaAAA	GACGTTTCCT	GGCAAACCGC	CCGGCAAAAa	gtcggtatgg
251	tctttcaaaag	taacgAactg	Tttgcccaca	tgaccgtcat	cgAaaacatc
301	ttcttAggcC	CGGTAAagga	aCAAAAcCgc	gaccgtgccc	aagcaGAGGC
351	gCAAGCCGGC	AAactGttgg	aacgcgTCGG	actgctAGAC	CGCAAAAACG
401	CCTATCCGCG	CGAactTTCC	GGCGGTCAGA	AACAGCGCAT	CGCCATTGTC
451	CGCGCCCTGT	GCCTGAATCC	GGAAGTCATC	CTGCTGGACG	AAATCACCGC
501	CGCACTTGAC	CCCGAAATGG	TGCGCGAAGT	CTTGGAAGTG	GTTTGGAAAC
551	TCGCCCCGCG	AGGGATGAGT	ATGCTCATCG	TAACCCACGA	AATGGGGTTC
601	GCACGCAAAG	TTGCCGACCG	CATCGTCTTT	ATGGACAAAG	GCGGCATCGT
651	CGAATCGTCC	GACCCGAAA	CCTTTTTTTC	CGACCAAAA	AGCGAACGCG
701	CCCGCAATT	TCTGGCAGGT	ATGGACTACT	GA	

This corresponds to the amino acid sequence <SEQ ID 446; ORF 122.ng>:

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 447>:

m122.seq

This corresponds to the amino acid sequence <SEQ ID 448; ORF 122>:

m122.per

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 122 shows 47.2% identity over a 246 aa overlap with a predicted ORF (ORF 122.ng) from *N. gonorrhoeae*:

m122/g122

	10	20	30	40	50	60
m122.pep	VVMIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQI					
	: : : : : : : : : : : : : : : : : : : : :					
g122	MALLSIRKLHKQYGSVTAIQSLDLLEKGEVIVLLGPSGCGKSTLLRCVNGLEPHQGGSII					
	10	20	30	40	50	60
	70	80	90	100	110	120
m122.pep	EFDNERPLKIDFSKKPSKHDILALRRKXSMVFQQYNLFPHKTALENVMEGPVAVQGKPAA					
	: : : : : : : : : : : :					
g122	VMDGVGEFGKDVSWQTA-----RQKVGVMVFQSNELFAHMTVIENIFLGPVKEQNRDRA					
	70	80	90	100	110	
	130	140	150	160	170	180
m122.pep	QAREEALKLLEKVGLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL					
	: : : : : : : : : : : : : : : :					
g122	EAEAQAGKLLERVGLLDRKNAYPRELSGGQKQRIAIVRALCLNPEVILLDEITAAALPEM					
	120	130	140	150	160	170
	190	200	210	220	230	240
m122.pep	VQDVLDXMKELAQEGWTMVVVTHEIKFALEVATTXVVMDXGVIVEQSPQDLFDHPKHER					
	: : : : : : : : : : : : : : :					
g122	VREVLVLEVLAREGMSMLIVTHEMGFARKVADRIVFMDKGGIVESSDPETFFSAPKSE					

350

	180	190	200	210	220	230
	250					
m122.pep	TRRFLSQIQSTKIX					
	: : :					
g122	ARQFLAGMDYX					
	240					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 449>:

```

a122.seq
1   GTTGTCATGA TTAAATCCG CAATATCCAT AAGACCTTCG GCAAAAATAC
51  CATTTTGCGC GGCATCAATT TGGATGTGTG CAAAGGGCAG GTGGTCGTCA
101 TCCTCGGGCC TTCCGGCTCA GGCAAAACGA CGTTTCTGCG ATGCCTAAAC
151 GCGTTGGAAG TGCCGAAGA CGGACAAATC GAGTTCGACA ACGAGCGACC
201 GCTGAAAATC GATTTTCTA AAAAACCAAG CAAACACGAT ATTTTGGCAC
251 TGCGCCGCAA ATCAGGCATG GTGTTTCAAC AATACAACCT CTTTCCGCAC
301 AAAACCGCCT TGGAAAACGT GATGGAAGGA CCGGTTGCCG TACAGGGCAA
351 GCCTGCCGCC CAAGCGCGCG AAGAGGCTCT GAAACTGCTG GAAAAAGTCG
401 GCTTGGGCGA CAAAGTGGAT TTGTATCCCT ACCAGCTTTC CGGCGGTCAG
451 CAGCAGCGCG TCGGCATTGC CCGAGCATTG GCGATTGAGC CCGAGCTGAT
501 GTTGTTTGAC GAACCCACTT CCGCGCTTGA CCCCAGATTG GTGCAAGACG
551 TGTGTAACGC CATGAAGGAA TTGGCGCGGG AAGGTTGGAC GATGGTCGTC
601 GTTACCCACG AAATCAAGTT CGCGCTGGAA GTTGCCACGA CCGTTGTCGT
651 GATGGACGGC GCGGTTATCG TAGAGCAGGG CAGCCCGAAA GAGTTGTTCT
701 ACCACCCCAA ACACGAACGG ACGCGGAGAT TTTTAAGCCA AATCCAATCT
751 ACCAAGATTG GA

```

This corresponds to the amino acid sequence <SEQ ID 450; ORF 122.a>:

```

a122.pep
1   VVMIKIRNIH KTFGKNTILR GINLDVCKGQ VVVLGPSPGS GKTTFRLCLN
51  ALEMPEDGQI EFDNERPLKI DFSKKPSKHD ILALRRKSGM VFQQYNLFPH
101 KTALENVMEG PVAVQGKPAA QAREEALKLL EKVLGDKVD LYPYQLSGGQ
151 QQRVGIARAL AIQPELMLFD EPTSALDPEL VQDVNLAMKE LAREGWTMVV
201 VTHEIKFALE VATTVVVMDG GVIVEQGSPPK ELFDHPKHER TRRFLSQIQS
251 TKI*

```

m122/a122 96.0% identity in 253 aa overlap

	10	20	30	40	50	60
m122.pep	VVMIKIRNIHKTFGENTILRGIDLVDCKGQVVVLGPSPGSGKTTFLRCLNALEMPEDGQI					
	: : : : : : : : : : :					
a122	VVMIKIRNIHKTFGKNTILRGINLDVCKGQVVVLGPSPGSGKTTFLRCLNALEMPEDGQI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m122.pep	EFDNERPLKIDFSKKPSKHDILALRRKSMVFQQYNLFPHKTALENVMEGPPVAVQGKPAA					
	: : : : : : : : : :					
a122	EFDNERPLKIDFSKKPSKHDILALRRKSGMVVFQQYNLFPHKTALENVMEGPPVAVQGKPAA					
	70	80	90	100	110	120
	130	140	150	160	170	180
m122.pep	QAREEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL					
	: : : : : : : : : :					
a122	QAREEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSALDPEL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m122.pep	VQDVLDMKELAQEGWTMVVVTHEIKFALEVATTXVMDXGVIVEQGSPPQDLFDHPKHER					
	: : : : : : : : : :					
a122	VQDVNLAMKELAREGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPPKELFDHPKHER					
	190	200	210	220	230	240
	250					
m122.pep	TRRFLSQIQSTKIX					

351

a122 |||||
 TRRFSLQIQSTKIX
 250

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 451>:

g122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACCATTTT
51  GCGCGGCATC GATTGGGATG TGGGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCCGG CTCGGGTAAA ACAACATTTC TGCCTGCCTT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGCGC GGCCGTTACG
201 CATTGATTTT TCCAAAAAAA CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAAGTCCGG AATGGTATTC CAACAATACA ACCTCTTCCC GCATAAAACC
301 GTGTTGGAAA ACGTGATGGA AGGGCCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTTTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGATAAAGT GGATTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGTGTCGGTA TCGCCCGCGC ACTGGCGATT CAGCCTGAAT TGATGCTGTT
501 TGACGAACCC ACTTCCGCGC TGGACCCCGA GTTGGTGCAA GACGTGTTGG
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCACGCT GGAAGTTGCC ACGAACGTCG TCGTGATGGA
651 CGGCGGCGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701 TCAAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTGCCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 452; ORF 122-1.ng>:

g122-1.pep

```

1  MIKIRNIHKT FGENTILRGI DLDVGKGQVV VILGPSGSGK TTFLRCLNAL
51  EMPEDGQIEF DNARPLRIDF SKKTSKHDIL ALRRKSGMVF QQYNLFPHKT
101 VLENVMEGPV AVQGKPAQA REEALKLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDMKELA REGWTMVVVT
201 HEIKFTLEVA TNVVMDGGV IVEQGSPEL FDHLKHERTR RFLSQIQSAK
251 I*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 453>:

m122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTTGGCGAAA ACACATTTT
51  GCGCGGCATC GATTGGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCCGG CTCAGGCAAA ACGACGTTTC TGCATGCCTT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACTGCGCC
251 GCAATCAGG CATGGTGTTC CAACAATACA ACCTCTTCCC GCACAAAACC
301 GCCTTGGAAA ACGTAATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGACAAAGT GGATTGTAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCCGCG ATTGGCGATT CAGCCTGAAC TGATGCTGTT
501 TGACGAACCG ACTTCCGCGC TCGATCCTGA ATTGGTGCAA GATGTTTGG
551 ATACCATGAA GGAATTGGCG CAAGAAGGCT GGACCATGGT TGTCGTTACG
601 CATGAAATCA AGTTCGCCTT AGAAGTGGCA ACCACCGTCG TCGTGATGGA
651 CGGCGGCGTT ATTGTGGAAC AAGGCAGCCC GCAAGATTTG TTCGACCACC
701 CCAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 454; ORF 122-1>:

m122-1.pep

```

1  MIKIRNIHKT FGENTILRGI DLDVCKGQVV VILGPSGSGK TTFLRCLNAL
51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAQA REEALKLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLDTMKELA QEGWTMVVVT
201 HEIKFALEVA TTVVMDGGV IVEQGSQDL FDHPKHERTR RFLSQIQSTK
251 I*

```

m122-1/g122-1 94.8% identity in 251 aa overlap

```

          10      20      30      40      50      60
m122-1.pep  MIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF
          |||||
g122-1      MIKIRNIHKTFGENTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF

```

352

	10	20	30	40	50	60
m122-1.pep	70	80	90	100	110	120
	DNERPLKIDFSKKPSKHDILALRRKSGVMVFQQYNLFPHKTALENVMEGPPVAVQGKPAAQA					
g122-1	DNERPLKIDFSKKPSKHDILALRRKSGVMVFQQYNLFPHKTVLENVMEGPPVAVQGKPAAQA					
	70	80	90	100	110	120
m122-1.pep	130	140	150	160	170	180
	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					
g122-1	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					
	130	140	150	160	170	180
m122-1.pep	190	200	210	220	230	240
	DVLDTMKELAQEGWTMVVVTHEIKFALEVATTVVVMDGGVIVEQGSPODLFDHPKHHERTR					
g122-1	DVLDTMKELAQEGWTMVVVTHEIKFTLEVATNVVMDGGVIVEQGSPODLFDHLKHHERTR					
	190	200	210	220	230	240
m122-1.pep	250					
	RFLSQIQSTKIX					
g122-1	RFLSQIQSAKIX					
	250					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 455>:

a122-1.seq

```

1  ATGATTAAAA TCCGCAATAT CCATAAGACC TTCGGCAAAA ATACCATTTT
51  GCGCGGCATC AATTGGGATG TGTGCAAAGG GCAGGTGGTC GTCATCCTCG
101 GGCCTTCGGG CTCAGGCAAA ACGACGTTTC TCGGATGCCT AAACGCGTTG
151 GAAATGCCCG AAGACGGACA AATCGAGTTC GACAACGAGC GACCGCTGAA
201 AATCGATTTT TCTAAAAAAC CAAGCAAACA CGATATTTTG GCACGCGGCC
251 GCAAAATCAGG CATGGTGTTC CAACAATACA ACCTCTTTCC GCACAAAACC
301 GCCTTGGAAA ACGTGATGGA AGGACCGGTT GCCGTACAGG GCAAGCCTGC
351 CGCCCAAGCG CGCGAAGAGG CTCTGAAACT GCTGGAAAAA GTCGGCTTGG
401 GCGACAAAGT GGATTGTGAT CCCTACCAGC TTTCCGGCGG TCAGCAGCAG
451 CGCGTCGGCA TTGCCCGAGC ATTGGCGATT CAGCCCGAGC TGATGTTGTT
501 TGACGAACCC ACTTCCGCGC TTGACCCCGA GTTGGTGCAA GACGTGTTGA
551 ACGCCATGAA GGAATTGGCG CGGGAAGGTT GGACGATGGT CGTCGTTACC
601 CACGAAATCA AGTTCGCGCT GGAAGTTGCC ACGACCGTTG TCGTGATGGA
651 CGGCGCGGTT ATCGTAGAGC AGGGCAGCCC GAAAGAGTTG TTCGACCACC
701 CCAAACACGA ACGGACGCGG AGATTTTAA GCCAAATCCA ATCTACCAAG
751 ATTTGA

```

This corresponds to the amino acid sequence <SEQ ID 456; ORF 122-1.a>:

a122-1.pep

```

1  MIKIRNIHKT FGKNTILRGI NLDVCKGQVV VILGPSGSGK TFLRCLNAL
51  EMPEDGQIEF DNERPLKIDF SKKPSKHDIL ALRRKSGMVF QQYNLFPHKT
101 ALENVMEGPV AVQGKPAAQA REEALKLLEK VGLGDKVDLY PYQLSGGQQQ
151 RVGIARALAI QPELMLFDEP TSALDPELVQ DVLNAMKELA REGWTMVVVT
201 HEIKFALEVA TTVVVMDDGGV IVEQGSPEL FHPKHHERTR RFLSQIQSTK
251 I*

```

a122-1/m122-1 97.2% identity in 251 aa overlap

	10	20	30	40	50	60
a122-1.pep	MIKIRNIHKTFGKNTILRGINLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
m122-1	MIKIRNIHKTFGKNTILRGIDLDVCKGQVVVILGPSGSGKTTFLRCLNALEMPEDGQIEF					
	10	20	30	40	50	60
a122-1.pep	70	80	90	100	110	120
	DNERPLKIDFSKKPSKHDILALRRKSGVMVFQQYNLFPHKTALENVMEGPPVAVQGKPAAQA					
m122-1	DNERPLKIDFSKKPSKHDILALRRKSGVMVFQQYNLFPHKTALENVMEGPPVAVQGKPAAQA					
	70	80	90	100	110	120
a122-1.pep	130	140	150	160	170	180
	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					
m122-1	REEALKLLEKVLGDKVDLYPYQLSGGQQQRVGIARALAIQPELMLFDEPTSAIDPELVQ					

g125.seq

This corresponds to the amino acid sequence <SEQ ID 458; ORF 125.ng>:

m125.seq

1	ATGTCGGGCA	ATGCCCTCTC	TCCTTCATCT	TCCTCCGCCA	TCGGGGTGTAT
51	TTGGTTCCGC	GCGGCGGTAT	CGATTGCCGA	AATCAGCACG	GGTACGCTGC
101	TTGCGCCTTT	GGGCTGGCAG	CGCGGTCTGG	CGGCTCTACT	TTTGGGTTCAT
151	GCCGCTCGGC	GCGCGCTGTT	TTTTCGCGCG	GCGTATATCG	GCGCAGTCGAC
201	CGGACGACAG	TCCGATGGAAA	GCGTGGCGCCT	GTGTTCTCGC	AAACCGCGTT
251	CAGTGCTGTT	TTCCGTGGCG	AATATGCTGC	AACTGGCCGG	CTGGACGCGG
301	GTGATGATTT	ACGCCCGCGC	AACGGTCAGC	TCCGCTTTGG	GCAAAGTGTT
351	GTGGGACGGC	GAATCTTTTG	TCTGTGTGGC	ATTGGCAAAC	GGCGCGCTGA
401	TTGTGCTGTG	GCTGTTTTTC	GGCGCACGCA	AAACGGCCGG	GCTGAAAACC
451	GTTTCGATGC	TGCTGATGCT	TTGGCGGTTT	CTGTGGCTGA	GTGCCGAAGT
501	CTTTTCCACG	GCGAGCAGCA	CCGCCGCACA	GGTTTCAGAC	GGCATGAGTT
551	TCGGAACGGC	AGTCGAGCTG	TCCGCCGTGA	TGCCGCTTTC	CTGGCTGCCG
601	CTTGCCGCCG	ACTACACGCG	CCACGCGCGC	CGCCCGTTTG	CGGCAACCCT
651	GACGGCAACG	CTCGCCTACA	CGCTGACCGG	CTGCTGGATG	TATGCCTTGG
701	GTTTGGCAGC	GGCGTGTGTT	ACCGGAGAAA	CCGACGTGGC	AAAAATCCCT
751	CTGGGCGCAr	GTTTGTGTGC	GGCAGGCATT	TTGGCGGTGC	TCTCTCTCCG
801	CGTTACCACA	ACGTTTCTCG	ATGCCTATTCT	CGCCGGCGCG	AGTGCGAACG

354

851 ACATTTCGCG GCGTTTTGCG GAAACACCCG TCGCTGTCrG CGTTACCCTG
 901 ATCGGCACGG TACTTGCCGT CATGCTGCCC GTTACCGAAT ATGAAAACCT
 951 CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCGGgC GGTTTTGATT
 1001 GCCGACTTTT TCGTCTTGAA ACGGCGTGA

This corresponds to the amino acid sequence <SEQ ID 460; ORF 125>:

m125.pep

1 MSGNASSPSS SSAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH
 51 AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFSA NMLQLAGWTA
 101 VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT
 151 VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP
 201 LAADYTRHAR RPFAATLTAT LAYTLTGCWM YALGLAAALF TGETDVAKIL
 251 LGAXLGAAGI LAVVLSTVTT TFLDAYSAGA SANNISARFA ETPVAVXVTL
 301 IGTVLAVMLP VTEYENFLLL IGSVFAPMAG GFDCRLFRLE TA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 125 shows 92.1% identity over a 343 aa overlap with a predicted ORF (ORF 125.ng) from *N. gonorrhoeae*:

m125/g125

	10	20	30	40	50	60
m125.pep	MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA					
	: :					
g125	MSGNASSPSSSSAIGLVWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m125.pep	AYIGALTGRSSMESVRLSFGKRGSVLFVANMLQLAGWTAVMIYAGATVSSALGKVLWDG					
	: :					
g125	AYIGALTGRSSMESVRLSFGKCGSVLFVANMLQLAGWTAVMIYVAGATVSSALGKVLWDG					
	70	80	90	100	110	120
	130	140	150	160	170	179
m125.pep	ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQ-VS					
	: :					
g125	ESFVWWALANGALIVLWLVFGARRTGGLKTVSMLLMLLAVLWLSVEVFASSGTNAAPAVS					
	130	140	150	160	170	180
	180	190	200	210	220	230
m125.pep	DGMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAAL					
	: :					
g125	DGMTFGTAVELSAVMPLSWLPLAADYTRQARRPFAATLTATLAYTLTGCWMYALGLAAAL					
	190	200	210	220	230	240
	240	250	260	270	280	290
m125.pep	FTGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVT					
	: :					
g125	FTGETDVAKILLGAGLGITGILAVVLSTVTTTFLDYSAGASANNISARFAEIPVAVGVT					
	250	260	270	280	290	300
	300	310	320	330	340	
m125.pep	LIGTVLAVMLPVTEYENFLLIGSVFAPMAGGFDCRLFRLETAX					
	: :					
g125	LIRTVLAVMLPVTEYKNFLLIRSVFGPMAGGFDCRLFCLKTAX					
	310	320	330	340		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 461>:

a125.seq

1 ATGTCGGGCA ATGCCTCCTC TCCTTCATCT TCCGCCGCCA TCGGGCTGAT
 51 TTGGTTCGGC GCGGCGGTAT CGATTGCCGA AATCAGCAGC GGTACACTGC
 101 TTGCGCCTTT GGGCTGGCAG CGCGGTCTGG CGGCTCTGCT TTTGGGTCAT

355

```

151 GCCGTCGGCG GCGCGCTGTT TTTTGGCGCG GCGTATATCG GCGCACTGAC
201 CGGACGCAGC TCGATGGAAA GCGTGCGCCT GTCGTTCCGC AAACGCGGTT
251 CAGTGCTGTT TTCCGTGGCG AATATGCTGC AACTGGCCGG CTGGACGGCG
301 GTGATGATT ACGCCGGCGC AACGGTCAGC TCCGCTTTGG GCAAAGTGTT
351 GTGGGACGGC GAATCTTTTG TCTGGTGGGC ATTGGCAAAC GGCGCGCTGA
401 TTGTGCTGTG GCTGGTTTTT GCGGCACGCA AAACAGCGCG GCTGAAAACC
451 GTTTCGATGC TGCTGATGCT GTTGGCGGTT CTGTGGCTGA GTGCCGAAGT
501 CTTTCCACG GCAGGCAGCA CCGCCGCACA GGTTCAGAC GGCATGAGTT
551 TCGGAACGGC AGTCGAGCTG TCCGCCGTGA TGCCGCTTTC TTGGCTGCCG
601 CTGGCCGCGC ACTACACGCG CCACGCGCGC CGCCCGTTTG CGGCAACCCT
651 GACGGCAACG CTCGCCTACA CGCTGACCGG CTGCTGATG TATGCCTTGG
701 GTTTGGCAGC GCGGTTGTTT ACCGGAGAAA CCGACGTGGC AAAAACTCTG
751 CTGGGCGCAG GTTTGGGTGC GGCAGGCATT TTGGCGGTGC TCCTGTCGAC
801 CGTTACCACC ACTTTTCTCG ATGCCTACTC CGCCGCGGTA AGTGCCAACA
851 ATATTTCCGC CAAACTTTCG GAAATACCCA TCGCCGTTGC CGTCGCCGTT
901 GTCGGCACAC TGCTTGCCGT CCTCTGCCC GTTACCGAAT ATGAAAATT
951 CCTGCTGCTT ATCGGCTCGG TATTTGCGCC GATGGCG.GC GGTTTTGATT
1001 GCCGACTTTT TCGTCTTGAA ACGGCGTGA

```

This corresponds to the amino acid sequence <SEQ ID 462; ORF 125.a>:

```

a125.pep
1  MSGNASSPSS SAAIGLIWFG AAVSIAEIST GTLLAPLGWQ RGLAALLLGH
51  AVGGALFFAA AYIGALTGRS SMESVRLSFG KRGSVLFVSA NMLQLAGWTA
101 VMIYAGATVS SALGKVLWDG ESFVWWALAN GALIVLWLVF GARKTGGLKT
151 VSMLLMLLAV LWLSAEVFST AGSTAAQVSD GMSFGTAVEL SAVMPLSWLP
201 LAADYTRHAR RFFAATLTAT LAYTLTGCWM YALGLAAALF TGETDVAKIL
251 LGAGLGAAGI LAVVLSTVTT TFLDAYSAGV SANNISAKLS EIPIAVAVAV
301 VGTLLAVLLP VTEYENFLLL IGSVFAPMAX GFDCRLFRLE TA*

```

m125/a125 95.6% identity in 342 aa overlap

```

          10      20      30      40      50      60
m125.pep  MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
a125      MSGNASSPSSSSAIGLIWFGAAVSIAEISTGTLLAPLGWQRGLAALLLGHAVGGALFFAA
          10      20      30      40      50      60

          70      80      90     100     110     120
m125.pep  AYIGALTGRSSMESVRLSFGKRGSVLFVSNMQLAGWTAVMIYAGATVSSALGKVLWDG
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
a125      AYIGALTGRSSMESVRLSFGKRGSVLFVSNMQLAGWTAVMIYAGATVSSALGKVLWDG
          70      80      90     100     110     120

          130     140     150     160     170     180
m125.pep  ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQVSD
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
a125      ESFVWWALANGALIVLWLVFGARKTGGLKTVSMLLMLLAVLWLSAEVFSTAGSTAAQVSD
          130     140     150     160     170     180

          190     200     210     220     230     240
m125.pep  GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAALF
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
a125      GMSFGTAVELSAVMPLSWLPLAADYTRHARRPFAATLTATLAYTLTGCWMYALGLAAALF
          190     200     210     220     230     240

          250     260     270     280     290     300
m125.pep  TGETDVAKILLGAXLGAAGILAVVLSTVTTTFLDAYSAGASANNISARFAETPVAVXVTL
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
a125      TGETDVAKILLGAGLGAAGILAVVLSTVTTTFLDAYSAGVSANNISAKLSEIPIAVAVAV
          250     260     270     280     290     300

          310     320     330     340
m125.pep  IGTVLAVMLPVTEYENFLLIGSVFAPMAGGFDCRLFRLETAX
          |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
a125      VGTLLAVLLPVTEYENFLLIGSVFAPMAXGFDCRLFRLETAX

```

356

310 320 330 340

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 463>:

```

g126.seq
1   AtgccgtcTG AAaccCcaaa ggcACGCCGC CGGCTTTCAG ACGGCATCGC
51  GTCCGACAAC CATACCAAAG AATCCATCAT GCTCACCctg tacggcGAAA
101 CTTTCCCTTC GCGGCTGCTg ctcggcacgG cggcctacCC GACCCCTGAA
151 ATCCTCAAAC AATCCGTCCG AACCGCCCGG CCCGCGATGA ttaccGTCTC
201 GCTGCGCCGC ACGGGATGCG GCGGCGAGGC GCACGGTCAG GGGTTTGGT
251 CGCTGCTTCA AGAAACCGGC GTTCCCGTCC TGCCGAACAC GGCAGGCTGC
301 CAAAGCGTGC AGGAAGCGGT AACGACGGCG CAAATGGCGC GCGAAGTGTT
351 TGAAACCGAT TGGATAAAAT TGGAACTCAT CGGCGACGAC GACACCTTGC
401 AGCCGGACGT GTTCCAATC GTCGAAGCGG CGGAAATCCT GATTAAAGAC
451 GGCTTCAAAG TGCTGCCTTA TTGCACCGAA GACCTGATTG CCTGCCGCCG
501 CCTGCTCGAT GCGGGCTGTC AGGCGTTGAT GCCGTGGGCG GCTCCCATCG
551 GCACGGGTTT GGGGGCGGTT CACGCCTATG CGCTCAAAT CCTGCGCGAA
601 CGCCTGCCCG ACACGCCGCT GATTATCGAC GCGGGCTTGG GTTTCCTTC
651 CCAAGCGGCA CAAGTGATGG AATGGGTTT TGACGGCGTA TTGTTAAACA
701 CCGCCGTTTC CCGCAGCGGC GACCCCGTCA ACATGGCGCG CGCCTTCGCA
751 CTCGCCGTCG AATCCGGACG GCTGGCATT GAAGCCGGGC CGGTCGAAGC
801 GCGAACCAAA GCCCAAGCCA GCACGCCGAC AGTCGGACAA CCGTTTGGC
851 ATTCGGCGGA ATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 464; ORF 126.ng>:

```

g126.pep
1   MPSETPKARR RLSDGIASDN HTKESIMLTL YGETFPSRL LGTAAYPTPE
51  ILKQSVRTAR PAMITVSLRR TGCGGEAHGQ GFWSLLQETG VPVLPNTAGC
101 QSVQEAVTTA QMAREVFETD WIKLELIGDD DTLQPDVFQL VEAABILIKD
151 GFKVLPHYCTE DLIACRRLLD AGCQALMPWA APIGTGLGAV HAYALKILRE
201 RLPDTPLIID AGLGLPSQAA QVMEWGFDFV LLNTAVSRSG DPVNMARAF
251 LAVESGRLAF EAGPVEARTK AQASTPTVGQ PFWHSAEY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 465>:

```

m126.seq (partial)
1   ..CACTATACAA AGGAACCCAT TATGCTCACC CTATACGGCG AAAC TTCC
51  CTCGCGGCTG CTGCTCGGCA CGGCTGCCTA CCCGACCCCC GAAATCCTCA
101 AACAAATCCAT CCAAACCGCC CAGCCTGCGA TGATTACCGT CTCGCTGCGC
151 CGCGCGGGAA GCGCGGGCGA GGCGCACGGT CAGGGGTTTT GGTGCTGCT
201 TCAAGAAACC GCGGTTCCCG TCCTGCCGAA CACGGCAGGC TGCCAAAGCG
251 TGCAGGAAGC GGTAAACGACG GCGCAAATGG CGCGCGAAGT GTTTGAAACC
301 GATTGGATAA AATTGGAAC CATCGGAGAT GACGACACCT TGCAGCCGGA
351 TGTGTTCCAG CTGTGCGAAG CGGCGGAAAT CCTGATTAAA GACGGCTTCA
401 AAGTGCTGCC TTATTGCACC GAAGACCTGA TTGCCTGCCG CCGCCTGCTC
451 GACGCGGGCT GTCAGGCGTT GATGCCGTGG GCGGCCCCGA TCGGCACGGG
501 TTTGGGCGCG GTTCACGCCT ACGCGTTGAA CGTCCTGCGC GAACGCCTGC
551 CCGACACGCC GCTGATTATC GACGCGGGCT TGGGTTTGCC CTCACAGGCG
601 GCACAAGTGA TGGAAATGGGG CTTTGACGGC GTGCTTTTGA ATACTGCCGT
651 TTCCCGCAGC GCGGATCCGG TCAATATGGC ACGCGCCTTC GCACTCGCCG
701 TCGAATCCGG ACGGCTGGCA TTTGAAGCCG GACCGGTCGA AGCACGCGAC
751 AAAGCGCAAG CCAGCACGCC GACAGTCGGA CAACCGTTTT GGCATTGCGC
801 GGAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 466; ORF 126>:

```

m126.pep (partial)
1   ..HYTKEPIMLT LYGETFPSRL LLGTAAYPTP EILKQSIQTA QPAMITVSLR
51  RAGSGGEAHG QGFWSLLQET GVPVLPNTAG CQSVQEAVTT AQMAREVFET

```

357

```

101 DWIKLELIGD DDTLQPDVFQ LVEAAEILIK DGFKVLPYCT EDLIACRRL
151 DAGCQALMPW AAPIGTGLGA VHAYALNVLR ERLPDTPLII DAGLGLPSQA
201 AQVMEWGF DG VLLNTAVSRG GDPVNMAF ALAVESGR LA FEAGPVEARD
251 KAQASTPTVG QPFWHSAEY*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 126 shows 95.9% identity over a 269 aa overlap with a predicted ORF (ORF 126.ng) from *N. gonorrhoeae*:

m126/g126

```

                                10      20      30      40
m126.pep                      HYTK EIMLTLYGETFPSRLLLGTAAYPTPEILKQSIQTAQ
                                ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      MPSETPKARRRLSDGIASDNHTKESIMLTLYGETFPSRLLLGTAAYPTPEILKQSVRTAR
                                10      20      30      40      50      60

                                50      60      70      80      90      100
m126.pep      PAMITVSLRRAGSGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREVFETD
                ||||| ||||| :| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      PAMITVSLRRRTGCGGEAHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREVFETD
                70      80      90      100      110      120

                                110      120      130      140      150      160
m126.pep      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      WIKLELIGDDDTLQPDVFQ LVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWA
                130      140      150      160      170      180

                                170      180      190      200      210      220
m126.pep      APIGTGLGAVHAYALNVLRERLPDTPLIIDAGLGLPSQA AQVMEWGF DG VLLNTAVSRSG
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      APIGTGLGAVHAYALKILRERLPDTPLIIDAGLGLPSQA AQVMEWGF DG VLLNTAVSRSG
                190      200      210      220      230      240

                                230      240      250      260      270
m126.pep      DPVNMAFALAVESGR LA FEAGPVEARD KAQASTPTVG QPFWHSAEYX
                ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
g126      DPVNMAFALAVESGR LA FEAGPVEART KAQASTPTVG QPFWHSAEYX
                250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 467>:

a126.seq

```

1   TTGTTAATCC ACTATACAAA GGAACCCATT ATGCTCACCC TGTACAGCGA
51  AACTTTCCTT TCGCGGCTGC TGCTCGGCAC AGCCGCCTAC CCGACCCCTG
101 AATCCTCAA ACAATCCGTC CGAACCGCCC GGCCCGCAT GATTACCGTC
151 TCGCTGCGCC GCGCGGGATG CGGCGGCGAG GCGCACGGTC AGGGGTTTTC
201 GTCGCTGCTT CAAGAAACCG GCGTTCCCGT CCTGCCGAAC ACGGCAGGCT
251 GCCAAAGCGT GCAGGAAGCG GTAACGACGG CGCAAATGGC GCGCGAAGTG
301 TTGAAACCG ATTGGATTAA ACTCGAACTC ATCGGCGACG ACGACACCTT
351 GCAGCCGAT GTGTTCCAAC TTGTCGAAGC GCGCGAAATC CTGATTAAAG
401 ACGGCTTCAA AGTGCTGCCT TATTGCACCG AAGACCTGAT TGCCTGCCGC
451 CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG ATGCGGTGGG CGGCCCCGAT
501 CGGCACGGGT TTGGGCGCGG TTCACGCCTA CGCGTTGAAC GTCCTGCGCG
551 AACGCCTGCC CGACACGCCG CTGATTATCG ACGCGGGCTT GGGTTTGCCC
601 TCACAGGCGG CACAAGTGAT GGAATGGGGC TTTGACGGCG TGCTTTTGAA
651 TACTGCCGTT TCCCGCAGCG GCGATCCGGT CAATATGGCA CGCGCCTTCG
701 CACTCGCCGT CGAATCCGGA CGGCTGGCAT TTGAAGCCGG ACCGGTCGAA
751 GCACGCGACA AAGCGCAAGC CAGCACGCCG ACAGTCGGAC AACCGTTTTG
801 GCATTGCGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 468; ORF 126.a>:

```
a126.pep
  1  LLIHYTKEPI MLTYLSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV
 51  SLRRAGCGGE AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV
101  FETDWIKLEL IGDODTLQPD VFQVEAAEI LIKDGFKVLP YCTEDLIACR
151  RLLDAGCQAL MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP
201  SQAAQVMEWG FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE
251  ARDKAQASTP TVGQPFWWSA EY*
```

m126/a126 98.1% identity in 269 aa overlap

```

      10      20      30      40      50
m126.pep  HYTKEPIMLTLYGETFPSRLLGTAAYPTPEILKQSIQTAQPAMITVSLRRAGSGGE
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126      LLIHYTKEPIMLTLYSETFPSRLLGTAAYPTPEILKQSVRTARPAMITVSLRRAGCGGE
          10      20      30      40      50      60

      60      70      80      90     100     110
m126.pep  AHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPD
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126      AHGQGFWSLLQETGVPVLPNTAGCQSVQEA VTTAQMAREVFETDWIKLELIGDDDTLQPD
          70      80      90     100     110     120

      120     130     140     150     160     170
m126.pep  VFQVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALN
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126      VFQVEAAEILIKDGFKVLPYCTEDLIACRRLLDAGCQALMPWAAPIGTGLGAVHAYALN
          130     140     150     160     170     180

      180     190     200     210     220     230
m126.pep  VLRERLPDTPLIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMARAFALAVESG
          |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
a126      VLRERLPDTPLIIDAGLGLPSQAAQVMEWGF DGVLLNTAVSRSGDPVNMARAFALAVESG
          190     200     210     220     230     240

      240     250     260     270
m126.pep  RLAFEAGPVEARDKAQASTPTVGQPFWWSAEYX
          |||:|||||:|||||:|||||:|||||:|||||:|||||
a126      RLAFEAGPVEARDKAQASTPTVGQPFWWSAEYX
          250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 469>:

```
g126-1.seq
  1  ATGCTCACCC TGTACGCGCA AACTTTCCTC TCGCGGCTGC TGCTCGGCAC
 51  GGCCGCTTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
101  GGCCCGCGAT GATTACCGTC TCGCTGCGCC GCACGGGATG CGGCGGCGAG
151  GCGCACGGTC AGGGGTTTGT GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201  CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251  CGCAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAAGTC
301  ATCGGCGACG ACGACACCTT GCAGCCGGAC GTGTTCCAAC TCGTCGAAGC
351  GCGGGAATC CTGATTAAAG ACGGCTTCAA AGTGCTGCCT TATTGCACCG
401  AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ATGCGGGCTG TCAGGCGTTG
451  ATGCCGTGGG CGGCTCCCAT CGGCACGGGT TTGGGGGCGG TTCACGCCTA
501  TGCCTCAAA ATCCTGCGCG AACGCCTGCC CGACACGCGG CTGATTATCG
551  ACGCGGGCTT GGGTTTGCTT TCCCAAGCGG CACAAGTGAT GGAATGGGGT
601  TTTGACGGCG TATTGTAAAC CACCGCCGTT TCCCGCAGCG GCGACCCCGT
651  CAACATGGCG CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701  TTGAAGCCGG GCCGGTCGAA GCGCGAACCA AAGCCCAAGC CAGCACGCGG
751  ACAGTCGGAC AACCGTTTTG GCATTCGGCG GAATATTGA
```

This corresponds to the amino acid sequence <SEQ ID 470; ORF 126-1.ng>:

```
g126-1.pep
  1  MLTYLGETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRTGCGGE
```

359

```

51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQLVAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALK ILRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARTKAQASTP
251 TVGQPFWHS A EY*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 471>:

```

m126-1.seq
1  ATGCTCACCC TATACGGCGA AACTTTCCCC TCGCGGCTGC TGCTCGGCAC
51  GGCTGCCTAC CCGACCCCGG AAATCCTCAA ACAATCCATC CAAACCGCCC
101 AGCCTGCGAT GATTACCGTC TCGCTGCGCC GCGCGGGAAG CGGCGGCGAG
151 GCGCACGGTC AGGGGTTTGT GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251 CGCAATATGGC GCGCGAAGTG TTTGAAACCG ATTGGATAAA ATTGGAATC
301 ATCGGAGATG ACGACACCTT GCAGCCGGAT GTGTCCAGC TTGTCGAAGC
351 GCGGGAATC CTGATTAAAG ACGGCTTCAA AGTGTGCTT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCGGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CCGCTTGAAC GTCCTGCGCG AACGCCTGCC CGACACGCGG CTGATTATCG
551 ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGTG TGCTTTTGAA TACTGCCGTT TCCCGCAGCG GCGATCCGGT
651 CAATATGGCA CGCGCTTCG CACTCGCCGT CGAATCCGGA CGGTGGCAT
701 TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCAGCCG
751 ACAGTCGGAC AACCGTTTGT GCATTGCGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 5; ORF 126-1>:

```

m126-1.pep
1  MLTLYGETFP SRLLLGTAAY PTPEILKQSI QTAQPAMITV SLRRAGSGGE
51  AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQLVAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGDPVNMA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWHS A EY*

```

m126-1/g126-1 96.9% identity in 262 aa overlap

	10	20	30	40	50	60
m126-1.pep	MLTLYGETF	PSRLLLGTAAYPT	PEILKQSIQTAQ	PAMITVSLRRAGSGGE	AHGQGFWSLL	
g126-1	MLTLYGETF	PSRLLLGTAAYPT	PEILKQSVRTAR	PAMITVSLRRTGCGGE	AHGQGFWSLL	
	10	20	30	40	50	60
m126-1.pep	QETGVPVLPNTAGCQSVQEA	VTTAQMAREV	FETDWIKLEL	IGDDDTLQPD	VFQLVAAEI	
g126-1	QETGVPVLPNTAGCQSVQEA	VTTAQMAREV	FETDWIKLEL	IGDDDTLQPD	VFQLVAAEI	
	70	80	90	100	110	120
m126-1.pep	LIKDGFKVLPYCTEDLIACR	RLLDAGCQAL	MPWAAPIGTGLGAVHAYALN	VLRERLPDTP		
g126-1	LIKDGFKVLPYCTEDLIACR	RLLDAGCQAL	MPWAAPIGTGLGAVHAYALK	ILRERLPDTP		
	130	140	150	160	170	180
m126-1.pep	LIIDAGLGLPSQAAQVMEWG	FDGVLLNTAVSRSGDPVNMA	RAFALAVESGRLAFEAGPVE			
g126-1	LIIDAGLGLPSQAAQVMEWG	FDGVLLNTAVSRSGDPVNMA	RAFALAVESGRLAFEAGPVE			
	190	200	210	220	230	240
m126-1.pep	ARDKAQASTPTVGQPFWHS	AEYX				
g126-1	ARTKAQASTPTVGQPFWHS	AEYX				
	250	260				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 473>:

```

a126-1.seq
1  ATGCTCACCC TGTACAGCGA AACTTTCCCT TCGCGGCTGC TGCTCGGCAC
51  AGCCGCTAC CCGACCCCTG AAATCCTCAA ACAATCCGTC CGAACCGCCC
101 GGCCCGCAT GATTACCGTC TCGCTGCGCC GCGCGGGATG CGGCGGCGAG

```

360

```

151 GCGCACGGTC AGGGGTTTGT GTCGCTGCTT CAAGAAACCG GCGTTCCCGT
201 CCTGCCGAAC ACGGCAGGCT GCCAAAGCGT GCAGGAAGCG GTAACGACGG
251 CGCAAATGGC GCGCGAAGTG TTTGAAACCG ATTGGATTAA ACTCGAACTC
301 ATCGGCGACG ACGACACCTT GCAGCCGGAT GTGTTCCAAC TTGTCGAAGC
351 GCGCGAAATC CTGATTAAAG ACGGCTTCAA AGTGTGCCT TATTGCACCG
401 AAGACCTGAT TGCCTGCCGC CGCCTGCTCG ACGCGGGCTG TCAGGCGTTG
451 ATGCCGTGGG CGGCCCCGAT CGGCACGGGT TTGGGCGCGG TTCACGCCTA
501 CGCGTTGAAC GTCCTGCCGC AACGCCTGCC CGACACGCCG CTGATTATCG
551 ACGCGGGCTT GGGTTTGCCC TCACAGGCGG CACAAGTGAT GGAATGGGGC
601 TTTGACGGCG TGCTTTTGAA TACTGCCGTT TCCGCGACGG GCGATCCGGT
651 CAATATGGCA CGCGCCTTCG CACTCGCCGT CGAATCCGGA CGGCTGGCAT
701 TTGAAGCCGG ACCGGTCGAA GCACGCGACA AAGCGCAAGC CAGCACGECG
751 ACAGTCGGAC AACCGTTTGT GCATTCGGCG GAATATTGA

```

This corresponds to the amino acid sequence <SEQ ID 474; ORF 126-1.a>:

a126-1.pep

```

1 MLTLYSETFP SRLLLGTAAY PTPEILKQSV RTARPAMITV SLRRAGCGGE
51 AHGQGFWSLL QETGVPVLPN TAGCQSVQEA VTTAQMAREV FETDWIKLEL
101 IGDDDTLQPD VFQVLEAAEI LIKDGFKVLP YCTEDLIACR RLLDAGCQAL
151 MPWAAPIGTG LGAVHAYALN VLRERLPDTP LIIDAGLGLP SQAAQVMEWG
201 FDGVLLNTAV SRSGBPVMNA RAFALAVESG RLAFEAGPVE ARDKAQASTP
251 TVGQPFWHSA EY*

```

a126-1/m126-1 98.1% identity in 262 aa overlap

	10	20	30	40	50	60
a126-1.pep	MLTLYSETFP	SRLLLGTAAY	PTPEILKQSV	RTARPAMITV	SLRRAGCGGE	AHGQGFWSLL
m126-1	MLTLYGETFP	SRLLLGTAAY	PTPEILKQSI	QTAQPAMITV	SLRRAGSGGE	AHGQGFWSLL
	10	20	30	40	50	60
	70	80	90	100	110	120
a126-1.pep	QETGVPVLP	NTAGCQSVQ	EAVTTAQMA	REVFETDWIK	LELIGDDDT	LQPDVFQVLEAAEI
m126-1	QETGVPVLP	NTAGCQSVQ	EAVTTAQMA	REVFETDWIK	LELIGDDDT	LQPDVFQVLEAAEI
	70	80	90	100	110	120
	130	140	150	160	170	180
a126-1.pep	LIKDGFKVLP	YCTEDLIAC	RRLLDAGCQ	ALMPWAAPIG	TGLGAVHAY	ALNVLRERLPDTP
m126-1	LIKDGFKVLP	YCTEDLIAC	RRLLDAGCQ	ALMPWAAPIG	TGLGAVHAY	ALNVLRERLPDTP
	130	140	150	160	170	180
	190	200	210	220	230	240
a126-1.pep	LIIDAGLGLP	SQAAQVMEW	GFDGVLLNT	AVSRSGBPVM	NARAFALAVE	SGRLAFEAGPVE
m126-1	LIIDAGLGLP	SQAAQVMEW	GFDGVLLNT	AVSRSGBPVM	NARAFALAVE	SGRLAFEAGPVE
	190	200	210	220	230	240
	250	260				
a126-1.pep	ARDKAQASTP	TVGQPFWHS	AEYX			
m126-1	ARDKAQASTP	TVGQPFWHS	AEYX			
	250	260				

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 475>:

g127.seq

```

1 ATGGAAATAT GGAATATGTT GAACACTTGG CCCGATGCCG TCCCGATACG
51 CGCGGAGGCG GCCGAATCCG TGGCGGCGGT CGCGGCTTTG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAGAC GGCATCCGGA TTTGCGCATC
151 GAAAGCAAGC GCGGTTTTTT GGTGCCAGC CGCAATATAA CGCTGCTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GGCAGAAATT CAAACGCTGG
251 CTTTGTCGAT GTTTGCGGTG GCGGCGGCGG TCGTCGTGGC GACAAAAGAA
301 CTGATTATGT GTCTGTCGGG CAGTATTTTA aggtctGCCA CCCAGCAATA
351 CTCGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401 ACATCAATCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
451 GTCGACAGC TTGCGGGAAC CACCGTTTCT TTCCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCCGCG ACAATATTTT GGGCGACTAT GTCATCCATA

```

361

```

551 CGGTCGAAAT CCCC GTTCCC ATCCATTGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
651 TCAGCGGTAT TTGGA AAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCG
701 CCGCCAGGCC GCGCGTTACC CGCGTACCGT ACGACGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCCTC CCCC GTTTC AAGCGGCTGG AAATCCAACA
801 GGCGGTTATG GACGAATTTT TGC GCGTACA ATACCGCCTG TTAAATCATC
851 CCGCCGgctc cgAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 476; ORF 127.ng>:

```

g127.pep
1  MEIWNMLNTW PDAVPIRAEA AESVA AVAL LLARALLNI HFRRHPDFGI
51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AA VVVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251 IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 477>:

```

m127.seq
1  ATGGAATAT GGAATATGTT GGACACTTGG CTCGGTGCCG TCCCGATACG
51  TCGCGAGGCG GTCGAATCCG TGGCGGCGGT TCGCGCTTTG CTGCTGGCGC
101 GCGCCCTTCT GTTGAATATC CACTTCAAAC GGCATCCGGA TTTCGGCATC
151 GAAAGCAAGC GCGCGTTTTT GGTGCGCAGC CGCAATATAA CGCTGCTTTT
201 GGTGCTGTTT TCGCTGGCAT TTATCTGGTC GCGCGCAAATC CAAACGCTGG
251 CTTTGTGAT GTTTGCGGTG GCGGCGGCGG TCGTCGTGGC GACGAAGGAA
301 CTGATTATGT GTCTGTCGGG CAGTATTTTA AGGTCGTCCA CCCAGCAATA
351 CTCGGTCGGC GACTATATCG AAATCAACGG CCTGCGCGGG CGCGTGGTCG
401 ACATCAACCT GTTGAACACG CTGATGATGC AGGTCGGTCC GAACCCCTTG
451 GTCGGACAGC TTGCGGGAAC CACCGTTTCT TTCCCAACA GCCTGTTGTT
501 GAGCCACCCC GTGCGCGCGC ACAATATTTT GGGCGACTAT GTCATCCATA
551 CGGTCGAAAT CCCC GTTCCC ATCCATTGG ATTCGGATGA AGCCGTATGC
601 CGTCTGAAAG CCGTACTCGA GCCCTTGTGC GCGCCCTACA TCCCCGCCAT
651 CCAACGgAT TTGGA AAAACG TGCAGGCGGA AAAACTGTTT ATCACGCCCG
701 CCGCCAGACC GCGCGTTACC CGCGTGCCGT ACGATGACAA GGCATACCGC
751 ATCATCGTCC GCTTCGCTTC CCCC GTTTC AAGCGGCTGG AAATCCAACA
801 GGCGGTTATG GACGAATTTT TGC GCGTACA ATACCGCCTG TTAAATCACC
851 CCGCCGGCTC CGAAACACTT TAA

```

This corresponds to the amino acid sequence <SEQ ID 478; ORF 127>:

```

m127.pep
1  MEIWNMLDTW LGAVPIRAEA VESVA AVAL LLARALLNI HFKRHPDFGI
51  ESKRRFLVAS RNITLLLVLF SLAFIWSAQI QTLALSMFAV AA VVVATKE
101 LIMCLSGSIL RSATQQYSVG DYIEINGLRG RVVDINLLNT LMMQVGNPL
151 VGQLAGTTVS FPNSLLLSHP VRRDNILGDY VIHTVEIPVP IHLDSDEAVC
201 RLKAVLEPLC APYIPAIQRY LENVQAEKLF ITPAARPRVT RVPYDDKAYR
251 IIVRFASPVS KRLEIQQAVM DEFLRVQYRL LNHPAGSETL *

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 127 shows 97.9% identity over a 290 aa overlap with a predicted ORF (ORF 127.ng)

from *N. gonorrhoeae*:

```

m127/g127
          10      20      30      40      50      60
m127.pep MEIWNMLDTWLGAVPIRAEAVESVA AVAL LLARALLNIHFKRHPDFGIESKRRFLVAS
|||||:| | |||||:| | |||||:| | |||||:| | |||||:| | |||||:| | |||||:| |
g127      MEIWNMLNTWPDVPIRAEAAESVA AVAL LLARALLNIHFRRHPDFGIESKRRFLVAS
          10      20      30      40      50      60

          70      80      90     100     110     120
m127.pep RNITLLLVLFSLAFIWSAQIQTALSMFAVA AAVVVATKELIMCLSGSILRSATQQYSVG
|||||:| | |||||:| | |||||:| | |||||:| | |||||:| | |||||:| | |||||:| |
g127      RNITLLLVLFSLAFIWSAQIQTALSMFAVA AAVVVATKELIMCLSGSILRSATQQYSVG

```

362

	70	80	90	100	110	120
	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY					
g127	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGQLAGTTVSFPNSLLLSHPVRRDNILGDY					
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT					
g127	VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRYLENVQAEKLFITPAARPRVT					
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLNHPAGSETLX					
g127	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLNHPAGSETLX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 479>:

```

a127.seq
1  ATGGAATAT  GGAATATGTT  GGACACTTGG  CTCGGTGCCG  TCCCGATACG
51  TCGCGAGGCG  GTCGAATCCG  TGGCGGTGGT  CGCGGCTTGT  CTGCTGGCGC
101 GCGCCCTTCT  GTTGAATATC  CACTTCAAAC  GGATCCGGA  TTTCGGCATC
151 GAAAGCAAGC  GCGCGTTTTT  GGTGCGCAGC  CGCAATATAA  CGCTGCTTTT
201 GGTGCTGTTT  TCGCTGGCAT  TTATCTGGTC  GCGCGAAATC  CAAACGCTGG
251 CTTTGTGCGAT  GTTTGCGGTG  GCGCGCGCGG  TCGTCGTGGC  GACGAAGGAA
301 CTGATTATGT  GTCTGTCGGG  CAGCATTTTA  AGGTCTGCCA  CCCAGCAATA
351 CTCGGTCGGC  GACTATATCG  AAATCAACGG  CCTGCGCGGG  CGCGTGGTCG
401 ACATCAACCT  GTTGAACACG  CTGATGATGC  AGGTGCGTCC  GAACCCCTTG
451 GTCGGACAGC  TTGCGGGAAC  CACCGTTTCT  TTCCCAACA  GCCTGTTGTT
501 GAGCCACCCC  GTGCGCGCGG  ACAATATTTT  GGGCGACTAC  GTCATCCATA
551 CGGTCGAAAT  CCCGGTTCCT  ATCCATTTGG  ATTCGGATGA  AGCCGTATGC
601 CGTCTGAAAG  CCGTACTCGA  GCCCTGTGTC  GCGCCCTACA  TCCCGGCCAT
651 CCAACGGCAT  TTGGAACACG  TGCAGGCGGA  AAACTGTTT  ATCAGCCCCG
701 CCGCCAAACC  GCGCGTTACC  CGCGTGCCGT  ACGATGACAA  GGCATACCGC
751 ATCATCGTCC  GCTTCGCCTC  CCCCCTTCA  AAGCGGCTGG  AAATCCAACA
801 GCGCGTTATG  GACGAATTTT  TGCAGTACA  ATACCGCCTG  TTAAATTACC
851 CCGCCGGCTC  CGAAACACTT  TAA

```

This corresponds to the amino acid sequence <SEQ ID 480; ORF 127.a>:

```

a127.pep
1  MEIWNMLDTW  LGAVPIRAEA  VESVAVVAAL  LLARALLLNI  HFKRHPDFGI
51  ESKRRFLVAS  RNITLLLVLF  SLAFIWSAQI  QTLALSMFAV  AAHVVVATKE
101 LIMCLSGSIL  RSATQQYSVG  DYIEINGLRG  RVVDINLLNT  LMMQVGNPL
151 VQLAGTTVS  FPNSLLLSHP  VRRDNILGDY  VIHTVEIPVP  IHLDSDEAVC
201 RLKAVLEPLC  APYIPAIQRH  LENVQAEKLF  ITPAAKPRVT  RVPYDDKAYR
251 IIVRFASPV  KRLEIQQAVM  DEFLRVQYRL  LNPAGSETL  *

```

m127/a127 98.6% identity in 290 aa overlap

	10	20	30	40	50	60
m127.pep	MEIWNMLDTWLGAVPIRAEAVESVAVVAALLARALLLNHFKRHPDFGIESKRRLVAS					
a127	MEIWNMLDTWLGAVPIRAEAVESVAVVAALLARALLLNHFKRHPDFGIESKRRLVAS					
	10	20	30	40	50	60
	70	80	90	100	110	120
m127.pep	RNITLLLVLFSLAFIWSAQIQTALSMFAVAAHVVVATKELIMCLSGSILRSATQQYSVG					
a127	RNITLLLVLFSLAFIWSAQIQTALSMFAVAAHVVVATKELIMCLSGSILRSATQQYSVG					
	70	80	90	100	110	120

363

	130	140	150	160	170	180
m127.pep	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGLAGTTVSFPNSLLLSHPVRRDNILGDY					
a127	DYIEINGLRGRVVDINLLNTLMMQVGNPLVGLAGTTVSFPNSLLLSHPVRRDNILGDY					
	130	140	150	160	170	180
	190	200	210	220	230	240
m127.pep	VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRXLENVQAEKLFITPAARPRVT					
a127	VIHTVEIPVPIHLDSDEAVCRLKAVLEPLCAPYIPAIQRHLENVQAEKLFITPAARPRVT					
	190	200	210	220	230	240
	250	260	270	280	290	
m127.pep	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLNHPAGSETLX					
a127	RVPYDDKAYRIIVRFASPVSKRLEIQQAVMDEFLRVQYRLNYPAGSETLX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 481>:

```

g128.seq
1  atgattgaca acgCactgct ccacttgggc gaagaaccCC GTTTaatca
51  aatccaaacc gaagACatca AACCcGCCGT CCAAACCGCC ATCGCCGAAG
101 CGCGCGGACA AATCGCCGCC GTCAAAGCGC AAACGCACAC CGGCTGGGCG
151 AACACCGTCG AGCGTCTGAC CGGCATCACC GAACGCGTCG GCAGGATTG
201 GGGCGTCGTG TCCCATCTCA ACTCCGTCGT CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCTGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAACTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTGCA ACGCTTTCCC CCGCACAAA AaccaagCTC GATCAGACC
401 TGCGCGATTT CGTATTGAGC GGCGCGGAAC TGCCGCCCGA ACGGCAGGCA
451 GAACTGGCAA AACTGCAAA CGAAGGCGCG CAACTTTCG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCCGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCCGCGCAAA GCGAAGGCAA AACAGGTTAC AAAATCGGCT TGCAGATTCC
651 GCACTACCTT GCCGTTATCC AATACGCCGG CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGTGCCA GCGAACTTTC AAACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCATTGAA
801 AACCgcaaaa cTGCTCGGCT TTAAAAATTA CGCCGAATTG TCGTGGCAA
851 CCAAATGGC GGACACGCCC GAACAGGTTT TAACTTCCT GCACGACCTC
901 GCCCGCGCG CCAAACCCTA CGCCGAAAA GACCTCGCG AAGTCAAAGC
951 CTTGCGCCG GAACACCTCG GTCTCGCCGA CCCGACGCG TGGGACTTGA
1001 GCTACGCCG CGAAAACTG CGCGAAGCCA AATACGATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTTCTGGCAG GCCTGTTGCG
1101 CCAAATCAAA AAACCTACG GCATCGGATT CGCCGAAAA ACCGTTCCCG
1151 TCTGGCACA AGACGTGCGC TATTTGAAT TGCAACAAA CGGCAAAACC
1201 ATCGGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGActaca AAGGCCGCG CCGCTTGCC GACGgcacGC
1301 TGCAACTGCC CACCGCTAC CTCGTCTGCA ACTTCGCCC GCCCGTCGGC
1351 GGCAAAGAAG CGCGTTAAG CCACGACGAA ATCCTCACCC TCTTCCACGA
1401 AacCGGCCAC GGAAGTGCAC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TGTCCGCGAT CAacggcgtA GAATGGGACG CGGTGGAAT GCCAGCCAG
1501 TTTATGGAAT ACTTCGTTG GGAATACAAT GTATTGGCAC AAATGTCCGC
1551 CCACGAAGAA AccgGCGAGC CCCTGCCGAA AGAACTCTT GACAAAATGC
1601 TcgCGCCAA AAACCTCCAG CGCGGTATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TCGATATGAT GATTACAGT GAAAGCGACG AATGCCGTCT
1701 GAAAACTGG CAGCAGGTTT TAGACAGCGT GCGCAAAGAA GTcGCCGTCA
1751 TCCAACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCacatctTC
1801 GCcggcGGCT ATTCCGAGG CTATTACAGC TACGCATGGG CCGAAGTCct
1851 cAGCACCGAT GCCTACGCCG CCTTGAAGA AAGcGACGac gtcGCCGCCA
1901 CAGGCAACG CTTCTGGCAA GAAAtccttg ccgtcggcgg ctCCCGCAGC
1951 gcgGCGGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCAG GCATAGACGC

```

2001 ACTGCTGCGC CAAagcggtT TCGACAACGC gGcttgA

This corresponds to the amino acid sequence <SEQ ID 482; ORF 128.ng>:

g128.pep

```

1 MIDNALLHLG EEPFRNQIQT EDIKPAVQTA IAEARGQIAA VKAQHTGTWA
51 NTVERLTGIT ERVGRINGVV SHLNSVVDTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFA TLSPAQKTKL DHDLRDFVLS GAELPPERQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTTY KIGLQIPHYL AVIQYAGNRE LREQIYRAYV TRASELSNDG
251 KFDNTANIDR TLENALKTAK LLGFKNYAEI SLATKMADTP EQVLNPLHDL
301 ARRAKPYAEK DLAEVKAFAR EHLGLADPQP WDSYAGEKL REAKYAFSET
351 EVKKYFPVVK VLAGLFAQIK KLYGIGFAEK TVPVWHKDVR YFELQQNGKT
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFA DGTLLQLPTAY LVCNFAPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGEPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDDMIYS ESDECRLKNW QQVLDSVRKE VAVIQPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSTD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR QSGFDNAA*
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 483>:

m128.seq (partial)

```

1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51 AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AACCGCACAC CGGCTGGGCA
151 AACACTGTCT AACCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTTG
201 GGGCGTGGTG TCGCACCTCA ACTGCGTCGC CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACCCAC
1 TACGCCAGCG AAAAAGTGGC CGAAGCCAAA TACGCGTTCA GCGAAACCGA
51 WGTCAAAAAA TAYTTCCCYG TCGGCAWGT ATTAACCGGA CTGTTCCGCC
101 AAmTCAAAAA ACTmTACGGC ATCGGATTTA CCGAAAAAAC yGTCCCCGTC
151 TGGCACAAAG ACGTGCCTTA TtTGAAATG CAACAAAACG GCGAaAmCCAT
201 AGGCGGCGTT TATATGGATT TGTACGCACG CGAAGGCAAA CGCGGCGGCG
251 CGTGGATGAA CGACTACAAA GGCGCGCGCC GTTTTTTCTA CGGCACGCTG
301 CAAYTGCCCA CCGCCTACCT CGTCTGCAAC TTCGCCCCAC CCGTCGCGCG
351 CAGGGAAGCC CGCyTAGGCC ACGACGAAAT CCTCATCTCT TCCACGAAA
401 CCGGACACGG GCTGCACCAC CTGCTTACCC AAGTGGACGA ACTGGGCGTA
451 TCCGGCATCA ACGGCGTAKA ATGGGACGCG GTCGAACTGC CCAGCCAGTT
501 TATGGAATAT TTCGTTTGGG AATACAATGT CTTGGCACAA mTGTCAGCCC
551 ACGAAGAAAC CGGcgTTCCC yTGCCGAAAG AACTCTTsGA CAAAwTGCTC
601 GCGGCCAAAA ACTTCCAAsG CGGCATGTTT yTsGTCCGGC AAwTGGAGTT
651 CGCCCTCTTT GATATGATGA TTTACAGCGA AGACGACGAA GGCCGTCTGA
701 AAAACTGGCA ACAGGTTTTA GACAGCGTGC GCAAAAAAGT CGCCGTCTATC
751 CAGCCGCCCG AATACAACCG CTTCCGCTTG AGCTTCGGCC ACATCTTCGC
801 AGCGGGCTAT TCCGACGCTn ATTACAGCTA CGCGTGGGCG GAAGTATTGA
851 GCGCGGACGC ATACGCCGCC TTTGAAGAAA GCGACGATGT CGCCGCCACA
901 GGCAAAACGCT TTGGCAGGA AATCCTCGCC GTCGGGGnAT CGCGCAGCGG
951 nGCAGAATCC TTCAAAGCCT TCCGCGGCCG CGAACCGAGC ATAGACGCAC
1001 TCTTGCGCCA CAGCGGTTTC GACAACGCGG TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 484; ORF 128>:

m128.pep (partial)

```

1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGTWA
51 NTVEPLTGIT ERVGRINGVV SHLNCVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSPAQKTKL NH

//

1 YASEKLREAK YAFSETXVKK YFPVGXVLNG LFAQXKKLYG IGFTEKTVPV
51 WHKDVRXYEL QQNGEXIGGV YMDLYAREGK RGGAWMNDYK GRRRFSDGTL
101 QLPTAYLVCN FAPPVGGREA RLSHDEILIL FHETGHGLHH LLTQVDELGV
151 SGINGVXWDA VELPSQFMEN FVWEYNVLAQ XSAHEETGVP LPKELXDKXL
201 AAKNFQXGMF XVRQXEFALF DMMIYSEDDE GRLLKNWQQVL DSVRKKVAVI
251 QPPEYNRFAL SFGHIFAGGY SAAXYSYAWA EVLSADAYAA FEESDDVAAT
```

301 GKRFWQEILA VGXSRGAES FKAFRGREPS IDALLRHSGF DNAV*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 128 shows 91.7% identity over a 475 aa overlap with a predicted ORF (ORF 128.ng) from *N. gonorrhoeae*:

m128/g128

	10	20	30	40	50	60
g128.pep	MIDNALLHLGEEPRFNQIQTEDIKPAVQTAIAEARGQIAAVKAQHTGTWANTVERLTGIT					
	:					
m128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGTWANTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
g128.pep	ERVGRIWGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA					
	:					
m128	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130	140	150	160	170	180
g128.pep	TLSPAQKTKLDHDLRDFVLGAELPPERQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	:					
m128	TLSPAQKTKLNH					
	130					
	//					
			340	350	360	
g128.pep	YAGEKLREAKYAFSETEVKKYFPVGKVLG					
	:					
m128	YASEKLREAKYAFSETXVKKYFPVGXVLNG					
			10	20	30	
	370	380	390	400	410	420
g128.pep	LFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWMNDYK					
	:					
m128	LFAQXKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWMNDYK					
	40	50	60	70	80	90
	430	440	450	460	470	480
g128.pep	GRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHLHLLLTQVDELGV					
	:					
m128	GRRRFSGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHLHLLLTQVDELGV					
	100	110	120	130	140	150
	490	500	510	520	530	540
g128.pep	SGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGEPLPKELFDKMLAAKNFQXGMF					
m128	SGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAAKNFQXGMF					
	160	170	180	190	200	210
	550	560	570	580	590	600
g128.pep	LVRQMEFALFDMMIYSESDCRLKNWQQVLDVSRKEVAVIQPPEYNRFANSFGHIFAGGY					
	:					
m128	XVRQXEFALFDMMIYSEDEGRLLKNWQQVLDVSRKKVAVIQPPEYNRFALSFGHIFAGGY					
	220	230	240	250	260	270
	610	620	630	640	650	660
g128.pep	SAGYYSYAWAEVLSTDAYAAFEESDDVAATGKRFWQEILAVGGSRAAESFKAFRGREPS					
	:					
m128	SAAXYSYAWAEVLSADAYAAFEESDDVAATGKRFWQEILAVGXSRSGAESFKAFRGREPS					
	280	290	300	310	320	330

366

```

          670      679
g128.pep  IDALLRQSGFDNAAX
          |||||:|||||:
m128      IDALLRHSGFDNAVX
          340

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 485>:

```

a128.seq
1  ATGACTGACA ACGCACTGCT CCATTGGGC GAAGAAGCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAACCGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTCTG AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTG
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CCGCCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAAATC AACCCAGATC
401 TCGCGGATTT CGTCCTCAGC GCGCGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAAAC CGAAGCGCGC CAACTTTCGG CCAAATTCTC
501 CCAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTC TTTGACGATG
551 CGCACCGCTT TGCCGGCATT CCGAAGACG CGTCGCCAT GTTTGCCGCT
601 GCGCGCAAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTATCC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAACTA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCGCGC CCAAACCTTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCGG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTCCG
1101 CCAATCAAAA AAACCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGCGGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGGCGG
1251 CGCGTGGATG AACGACTACA AAGGCCGCGC CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAGAAG CCCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGGAAT ATTTCTTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCAGCGCCAA AAATTTCCAA CGCGGAATGT TCCTCGTCCG CCAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAACTGG CAACAGGTTT TAGACAGCGT GCGCAAAGAA GTCGCCGTCG
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCCGCAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAACG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAAT CCTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA

```

This corresponds to the amino acid sequence <SEQ ID 486; ORF 128.a>:

```

a128.pep
1  MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51  NTVEPLTGIT ERVGRWGVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLSHAQKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEI SLATKMADTP EQVINFLHDL
301 ARRAKPYAEE DLAEVKAFAR ESLGLADLPQ WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKRRRFS DGTQLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME

```

551 FALFDMMIYS EDDEGRKLNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
 601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
 651 AAESFKAFRG REPSIDALLR HSGFDNAA*

m128/a128 66.0% identity in 677 aa overlap

	10	20	30	40	50	60
m128.pep	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAA	KAQHTGTWANTVEPLTGIT				
a128	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAA	KAQHTGTWANTVEPLTGIT				
	10	20	30	40	50	60
	70	80	90	100	110	120
m128.pep	ERVGRIWGVVSHLNCVADTPELRAVYNELMPEITVFFTEIGQDI	ELYNRFKTIKNSPEFD				
a128	ERVGRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDI	ELYNRFKTIKNSPEFD				
	70	80	90	100	110	120
	130					
m128.pep	TLSPAQKTKLNH-----					
a128	TLSHAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
m128.pep	-----					
a128	FDDAAPLAGIPEDALAMFAAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLRQEIYRAYV					
	190	200	210	220	230	240
m128.pep	-----					
a128	TRASELSDDGKFDNTANIDRTLLENALQTAKLLGFKNYAELSLATKMADTPEQVLNLFHDL					
	250	260	270	280	290	300
m128.pep	-----					
				140	150	
				YASEKLEAKYAFSETXVKKYFPVGX		
a128	ARRAKPYAEKDIAEVKAFARESGLADLQPWDLYAGEKLREAKYAFSETEVKKYFPVGK					
	310	320	330	340	350	360
m128.pep	160	170	180	190	200	210
	VLNGLFAQKKLYGIGFTEKTVPVWHKDVRYXELQQNGEXIGGVYMDLYAREGKRGGAWM					
a128	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
m128.pep	220	230	240	250	260	270
	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILILFHETGHGLHLLTQVD					
a128	NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHGLHLLTQVD					
	430	440	450	460	470	480
m128.pep	280	290	300	310	320	330
	ELGVSGINGVXWDAVELPSQFMENFVWEYNVLAQXSAHEETGVPLPKELXDKXLAANKFQ					
a128	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAANKFQ					
	490	500	510	520	530	540
m128.pep	340	350	360	370	380	390
	XGMFXVRQXEFALFDMMIYSEDDDEGRKLNWQQVLDSVRKKVAVIQPPEYNRFALSFGHIF					
a128	RGMFLVRQMEFALFDMMIYSEDDDEGRKLNWQQVLDSVRKEVAVVRPPEYNRFANSFGHIF					
	550	560	570	580	590	600

q128-1.seq (partial)

This corresponds to the amino acid sequence <SEQ ID 488; ORF 128-1.ng>:

g128-1.pep (partial)

1	MIDNALLHLG	EERFRNQIKT	EDIKPAVQTA	IAEARGQIAA	VKAQTHTGWA
51	NTVERLTGNT	ETRGVRIGVV	SHLSVVDPT	ELRAVYNELM	PEITVFFTEI
101	GDQIELYNR	KTKINSPFEA	KLQCGKTKL	DHDLRDFVLS	GADLPPERQA
151	ELAKLQTEGA	QLSAKFSQNV	LDATDAFGIY	FDDAAPLAGI	PEDALAMFAA
201	AAQSEKTYG	KIGLQIPIYL	AVIQYGNR	LRQIYRAYV	TRASELSNDG
251	KFDNTANIDR	TLENALKTAK	LFIQFVNAEL	SLATKMDATP	EQVLMFLNDL
301	ARRAKPYAEK	DLAEVKAFA	EHLGLADPQP	WDLSYAGEKL	REAKYAFSET
351	EVKKYPVGVK	VLAGLFAQIK	KLYGIGFAEK	TPVPVHKDVR	YFELQQNGKT
401	IGGVYMDLYA	REGKRGGAWM	KDYGRRRFAE	DGTLQLPTAY	LVCNFAPPVG
451	GKEARLSHDE	LITLPHETGH	GLHHLLTQVD	ELGVSGINGV	K

m128-1.seq

1 ATGACTGACA ACGCACTGCT CCATTTGGGC GAAGAACCCC GTTTTGATCA
51 AATCAAAACC GAAGACATCA AACC CGCCCT GCAAACCGCC ATCGCCGAAG
101 CGCGCGAACA AATCGCGGCC ATCAAAGCCC AAACGCACAC CGCGTGGGCA
151 AACACTGTGC AACCCCTGAC CGGCATCACC GAACGCGTCG CGGAGATTTG

```
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCGC CGACACGCCC GAACTGCGCG
251 CCGTCTATAA CGAACTGATG CCCGAAATCA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAATTCCCC
351 CGAATTCGAC ACCCTCTCCC CCGCACAAAA AACCAAATC AACCACGATC
401 TGC GCGATTT CGTCTCAGC GCGCGGAAC TGCCGCCGA ACAGCAGGCA
451 GAACTGGCAA AACTGCAAAC CGAAGGCGCG CAACTTTCG CCAAATTCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTTAC TTTAGCAGTG
551 CCGCACGCT TCGCGGCATT CCGAAGACG CGCTCGCCAT GTTTGCCGCC
601 GCGCGCAAA GCGAAAGCAA AACAGGCTAC AAAATCGGCT TGCAGATTCC
651 AACTACCTC GCGTCATCC AATACGCGA CAACCGCGAA CTGCGCGAAC
701 AAATCTACCG CGCTACGTT ACCCGCGCA GCGAACTTTC AGACGACGGC
751 AAATTCGACA ACACCGCAA CATCGACCGC ACGCTCGCAA ACGCCTGCA
801 AACCGCCAAA CTGCTCGGCT TCAAAAACTA CGCCGAATTG TCGTGGCAA
851 CCAAAATGGC GGACACGCCC GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCGCGCGCG CCAAACCCTA CGCCGAAAAA GACCTCGCG AAGTCAAAGC
951 CTTCGCGCG GAAAGCCTGA ACCTCGCGA TTTGCAACCG TGGGACTTGG
1001 GCTACGCCAG CGAAAACTG CGCGAAGCCA AATACGCGTT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAA GTATTAAACG GACTGTTGCG
1101 CCAAAATCAA AACTCTACG GCATCGGATT TACCGAAAA ACCGTCCCGG
1151 TCTGGCACA AGACGTGCGC TATTTTGAAT TGCAACAAA CGGCGAAACC
1201 ATAGGCGCGC TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGCGG
1251 CGCGTGGAT AACGACTACA AAGGCCGCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCGCCCC ACCCGTCGGC
1351 GGCAGGGAAG CCCGCTGAG CCACGACGAA ATCCTCATCC TCTTCCACGA
1401 AACCGGACAC GGGCTGCACC ACCTGCTTAC CCAAGTGGAC GAACTGGGCG
1451 TATCCGCGAT CAACGCGTA GAATGGGACG CGGTGCAACT GCCCAGCCAG
1501 TTTATGGAAT ATTTGTTTG GGAATACAA GTCTTGGCAC AAATGTCAGC
1551 CCACGAAGAA ACCGCGGTT CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCCGCCAA AACTTCCAA CGCGGCATGT TCCTCGTCCG GCAAATGGAG
1651 TTCGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAACTGG CAACAGGTT TAGACAGCGT GCGCAAAAA GTCCCGTCA
1751 TCCAGCCGCC CGAATACAA CGCTTCGCC TGAGCTTCGG CCACATCTTC
1801 GACGCGGCT ATTCCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCG CTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAACG CTTTGGCAG GAAATCCTCG CCGTCGCGG ATCGCGCAGC
1951 GCGCGAGAAT CCTTCAAAG CTTCCGCGG CGCGAACC GAATAGACGC
2001 ACTCTGCGC CACAGCGGTT TCGACAACGC GGTCTGA
```

This corresponds to the amino acid sequence <SEQ ID 490; ORF 128-1>:

m128-1.pep.

```
1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQHTGWAA
51 NTVEPLTGIT ERVGRIGVV SHLNSVADTP ELRAVYNELM PEITVFFTEI
101 GQDIELYNRF KTIKNSPEFD TLPQAKTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSESKTGY KIGLQIPHYL AVIQYADNRE LREQIYRAYV TRASELSDDG
251 KPDNTANIDR TLANALQTA LLGFKNYAEL SLATKMADTP EQVLNLFHDL
301 ARRAKPYAEK DLAEVKAFAR ESLNLADLQP WDLGYASEKL REAKYAFSET
351 EVKKYFPVGK VINGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTQLQPTAY LVCNFAPPVG
451 GREARLSHDE ILILFHETGH GLHLLTQVD ELGVSGINGV EWDAVELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLLKNW QQVLDSVRKK VAVIQPPEYN RFALSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAV*
```

m128-1/g128-1 94.5% identity in 491 aa overlap

```
10 20 30 40 50 60
g128-1.pep MIDNALLHLGEEPRFNQIKTEDIKPAVQTAIAEARGQIAAVKAQHTGWANTVERLTGIT
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128-1 MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQHTGWANTVEPLTGIT
10 20 30 40 50 60

70 80 90 100 110 120
g128-1.pep ERVGRIGVVSHLNSVVDTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFA
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m128-1 ERVGRIGVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD
70 80 90 100 110 120
```

370

	130	140	150	160	170	180
g128-1.pep	TLSPAQKTKLDHDLRDFVLSGAELPPERQAEALAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128-1	TLSPAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
	190	200	210	220	230	240
g128-1.pep	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYAGNRELREQIYRAYV					
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV					
	190	200	210	220	230	240
	250	260	270	280	290	300
g128-1.pep	TRASELSNDGKFDNTANIDRTLENALKTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
	250	260	270	280	290	300
	310	320	330	340	350	360
g128-1.pep	ARRAKPYAEKDLAEVKAFAREHLGLADPQPWDL SYAGEKLREAKYAFSETEVKKYFPVVGK					
m128-1	ARRAKPYAEKDLAEVKAFARESLNLADLPWDLGYASEKLREAKYAFSETEVKKYFPVVGK					
	310	320	330	340	350	360
	370	380	390	400	410	420
g128-1.pep	VLAGLFAQIKKLYGIGFAEKTVPVWHKDVRYFELQQNGKTIGGVYMDLYAREGKRGGAWM					
m128-1	VLNGLFAQIKKLYGIGFTEKTVPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM					
	370	380	390	400	410	420
	430	440	450	460	470	480
g128-1.pep	NDYKGRRRFADGTLQLPTAYLVCNFAPPVGGKEARLSHDEILTLFHETGHGLHHLLTQVD					
m128-1	NDYKGRRRFSDGTLQLPTAYLVCNFAPPVGGREARLSHDEILTLFHETGHGLHHLLTQVD					
	430	440	450	460	470	480
	490					
g128-1.pep	ELGVSGINGVK					
m128-1	ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ					
	490	500	510	520	530	540

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 491>:

a128-1.seq

```

1  ATGACTGACA ACGCACTGCT CCATTGGGCG GAAGAACCCC GTTTTGATCA
51  AATCAAAACC GAAGACATCA AACCCGCCCT GCAAAACGCC ATTGCCGAAG
101 CGCGCGAACA AATCGCCGCC ATCAAAGCCC AAACGCACAC CGGCTGGGCA
151 AACACTGTGC AACCCCTGAC CGGCATCACC GAACGCGTCG GCAGGATTGT
201 GGGCGTGGTG TCGCACCTCA ACTCCGTCAC CGACACGCCC GAACTGCGCG
251 CCGCTACAA TGAATTAATG CCCGAAATTA CCGTCTTCTT CACCGAAATC
301 GGACAAGACA TCGAGCTGTA CAACCGCTTC AAAACCATCA AAAACTCCCC
351 CGAGTTCGAC ACCCTCTCCC ACGCGCAAAA AACCAACTC AACCACGATC
401 TCGCGGATT TCGTCTCAGC GCGCGGGAAC TGCCGCCCGA ACAGCAGGCA
451 GAATTGGCAA AACTGCAAA CGAAGGCGCG CAACTTTCGG CCAAAATCTC
501 CCAAAACGTC CTAGACGCGA CCGACGCGTT CGGCATTAC TTTGACGATG
551 CCGCACCGCT TGCCGGCATT CCCGAAGACG CGCTCGCCAT GTTTGCCGCT
601 GCCGCGCAAA GCGAAGGCAA AACAGGCTAC AAAATCGGTT TGCAGATTCC
651 GCACTACCTC GCCGTACATC AATACGCCGA CAACCGCAAA CTGCGCGAAC
701 AAATCTACCG CGCCTACGTT ACCCGCGCCA GCGAGCTTTC AGACGACGGC
751 AAATTCGACA ACACCGCCAA CATCGACCGC ACGCTCGAAA ACGCCCTGCA
801 AACCGCAAAA CTGCTCGGCT TCAAAAATA CGCCGAATTG TCGCTGGCAA
851 CCAAAATGGC GGACACCCCG GAACAAGTTT TAAACTTCCT GCACGACCTC
901 GCCCGCCGCG CCAAAACCTA CGCCGAAAAA GACCTCGCCG AAGTCAAAGC
951 CTTGCGCCGC GAAAGCCTCG GCCTCGCCGA TTTGCAACCG TGGCACTTGG
1001 GCTACGCGCG CGAAAACTG CGCGAAGCCA AATACGCATT CAGCGAAACC
1051 GAAGTCAAAA AATACTTCCC CGTCGGCAAA GTATTAAACG GACTGTTTCG
1101 CCAATCAAAA AAATCTACG GCATCGGATT TACCGAAAAA ACCGTCCCCG
1151 TCTGGCACAA AGACGTGCGC TATTTGAAT TGCAACAAAA CGCGGAAACC
1201 ATAGCGCGCG TTTATATGGA TTTGTACGCA CGCGAAGGCA AACGCGCGCG

```



```
1251 CGCGTGGATG AACGACTACA AAGGCCGCCG CCGTTTTTCA GACGGCACGC
1301 TGCAACTGCC CACCGCCTAC CTCGTCTGCA ACTTCACCCC GCCCGTCGGC
1351 GGCAAGAAG CCGCTTGAG CCATGACGAA ATCCTCACCC TCTTCCACGA
1401 AACCGGACAC GGCCTGCACC ACCTGCTTAC CCAAGTCGAC GAACTGGGCG
1451 TATCCGGCAT CAACGGCGTA GAATGGGACG CAGTCGAACT GCCCAGTCAG
1501 TTTATGAAA ATTTTCGTTG GGAATACAAT GTCTTGGCGC AAATGTCCGC
1551 CCACGAAGAA ACCGGCGTTC CCCTGCCGAA AGAACTCTTC GACAAAATGC
1601 TCGCGCCCAA AAACCTCCAA CGCGGAATGT TCCTCGTCCG CCAAAATGGAG
1651 TCGGCCCTCT TTGATATGAT GATTACAGC GAAGACGACG AAGGCCGTCT
1701 GAAAAAAGTG CAACAGGTTT TAGACAGCGT GCGCAAGAA GTCGCCGTCT
1751 TCCGACCGCC CGAATACAAC CGCTTCGCCA ACAGCTTCGG CCACATCTTC
1801 GCAGGCGGCT ATTCGCGAGG CTATTACAGC TACGCGTGGG CGGAAGTATT
1851 GAGCGCGGAC GCATACGCCG CCTTTGAAGA AAGCGACGAT GTCGCCGCCA
1901 CAGGCAAAAG CTTTGGCAG GAAATCCTCG CCGTCGGCGG ATCGCGCAGC
1951 GCGGCAGAA CTTCAAAGC CTTCCGCGGA CGCGAACCGA GCATAGACGC
2001 ACTCTTGCGC CACAGCGGCT TCGACAACGC GGCTTGA
```

This corresponds to the amino acid sequence <SEQ ID 492; ORF 128-1.a>:

a128-1.pep

```
1 MTDNALLHLG EEPFRDQIKT EDIKPALQTA IAEAREQIAA IKAQTHTGWA
51 NTVEPLTGIT ERVGRIGVVV SHLNSVTDTP ELRAAYNELM PEITVFFTEI
101 QQDIELYNRF KTIKNSPEFD TLSHAQTKL NHDLRDFVLS GAELPPEQQA
151 ELAKLQTEGA QLSAKFSQNV LDATDAFGIY FDDAAPLAGI PEDALAMFAA
201 AAQSEGKTGY KIGLQIPHYL AVIQYADNRK LREQIYRAYV TRASELSDDG
251 KFDNTANIDR TLENALQTAK LLGFKNYAEL SLATKMADTP EQVLNFLHDL
301 ARRAKPYAEK DLAEVKAFAR ESLGLADLPQ WDLGYAGEKL REAKYAFSET
351 EVKKYFPVGK VLNGLFAQIK KLYGIGFTEK TVPVWHKDVR YFELQONGET
401 IGGVYMDLYA REGKRGGAWM NDYKGRRRFS DGTLLQLPTAY LVCNFTPPVG
451 GKEARLSHDE ILTLFHETGH GLHHLLTQVD ELGVSGINGV EWDDELPSQ
501 FMENFVWEYN VLAQMSAHEE TGVPLPKELF DKMLAAKNFQ RGMFLVRQME
551 FALFDMMIYS EDDEGRLEKNW QQVLDSVRKE VAVVRPPEYN RFANSFGHIF
601 AGGYSAGYYS YAWAEVLSAD AYAAFEESDD VAATGKRFWQ EILAVGGSRS
651 AAESFKAFRG REPSIDALLR HSGFDNAA*
```

m128-1/a128-1 97.8% identity in 677 aa overlap

	10	20	30	40	50	60
a128-1.pep	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT					
m128-1	MTDNALLHLGEEPRFDQIKTEDIKPALQTAIAEAREQIAAIIKAQTHTGWANTVEPLTGIT					
	10	20	30	40	50	60
	70	80	90	100	110	120
a128-1.pep	ERVGRIGVVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
m128-1	ERVGRIGVVVSHLNSVADTPELRAVYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFD					
	70	80	90	100	110	120
	130	140	150	160	170	180
a128-1.pep	TLSHAQTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
m128-1	TLSPAQTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIY					
	130	140	150	160	170	180
	190	200	210	220	230	240
a128-1.pep	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRKLREQIYRAYV					
m128-1	FDDAAPLAGIPEDALAMFAAAQSEKGTGYKIGLQIPHYLAVIQYADNRELREQIYRAYV					
	190	200	210	220	230	240
	250	260	270	280	290	300
a128-1.pep	TRASELSDDGKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
m128-1	TRASELSDDGKFDNTANIDRTLANALQTAKLLGFKNYAELSLATKMADTPEQVLNFLHDL					
	250	260	270	280	290	300
	310	320	330	340	350	360
a128-1.pep	ARRAKPYAEKDLAEVKAFARESLGLADLPQWDLGYAGEKLREAKYAFSETEVKKYFPVGK					

372

```

m128-1      ARRAKPYAEKDLAEVKAFARESLNLADLPWDLGYASEKLREAKYAFSETEVKKYFPVVGK
              310      320      330      340      350      360

              370      380      390      400      410      420
a128-1.pep  VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
              |||||
m128-1      VLNGLFAQIKKLYGIGFTEKTPVWHKDVRYFELQQNGETIGGVYMDLYAREGKRGGAWM
              370      380      390      400      410      420

              430      440      450      460      470      480
a128-1.pep  NDYKGRRRFS DGT LQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHLHLLTQVD
              |||||
m128-1      NDYKGRRRFS DGT LQLPTAYLVCNFTPPVGGKEARLSHDEILTLFHETGHLHLLTQVD
              430      440      450      460      470      480

              490      500      510      520      530      540
a128-1.pep  ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
              |||||
m128-1      ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ
              490      500      510      520      530      540

              550      560      570      580      590      600
a128-1.pep  RGMFLVRQMEFALFDMMIYSEDDGRLKNWQQVLDVRKEVAVVRPEYNRFANSFGHIF
              |||||
m128-1      RGMFLVRQMEFALFDMMIYSEDDGRLKNWQQVLDVRKKAQVIQPEYNRFANSFGHIF
              550      560      570      580      590      600

              610      620      630      640      650      660
a128-1.pep  AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFPWQEILAVGGSRSAAESFKAFRG
              |||||
m128-1      AGGYSAGYYSYAWAEVLSADAYAAFEESDDVAATGKRFPWQEILAVGGSRSAAESFKAFRG
              610      620      630      640      650      660

              670      679
a128-1.pep  REPSIDALLRHSGFDNAAX
              |||||
m128-1      REPSIDALLRHSGFDNAVX
              670

```

a128-1/ P44573

sp|P44573|OPDA_HAEIN OLIGOPEPTIDASE A >gi|1075082|pir||C64055 oligopeptidase A (prlC) homolog
- Haemophilus influenzae (strain Rd KW20)

>gi|1573174 (U32706) oligopeptidase A (prlC) [Haemophilus influenzae Rd] Length = 681

Score = 591 bits (1507), Expect = e-168

Identities = 309/677 (45%), Positives = 415/677 (60%), Gaps = 4/677 (0%)

```

Query: 4      NALLHLGEEPRFDQIKTEDIKPALQTXXXXXXXXXXXXXHTGTWANTVEPLTGITERV 63
              N LL++ P F QIK E I+PA++ H W N + PLT +R+
Sbjct: 5      NPLLNIQGLPPFSQIKPEHIRPAVEKLIQDCRNTIEQVLKQPHFTWENFILPLTETNDRL 64

Query: 64      GRIWGVVSHLNSVTDTPELRAAYNELMPEITVFFTEIGQDIELYNRFKTIKNSPEFDTLS 123
              R W VSHLSV ++ ELR AY +P ++ + T +GQ LYN + +KNS EF S
Sbjct: 65      NRAWSPVSHLNSVKNSTELREAYQTCLPLLSEYSTWVGQHKGLYNAYLALKNSAEFADYS 124

Query: 124     HAQKTKLNHDLRDFVLSGAELPPEQQAELAKLQTEGAQLSAKFSQNVLDATDAFGIYFDD 183
              AQK + + LRDF LSG L E+Q ++ ++L+++FS NVLDAT + ++
Sbjct: 125     IAQKKAIENSLRDFELSGIGLSEKQQRYSGEIVARLSELNSQFSNNVLDATMGWEKLIEN 184

Query: 184     AAPLAGIPEDALAMFAAAQSEGKTGYKIGLQIPHYLAVIQYADNRKLRQIYRAYVTRA 243
              A LAG+PE AL +A+S+G GY+ L+IP YL V+ Y +NR LRE++YRAY TRA
Sbjct: 185     EAELAGLPESALQAAQQAESKGLKGYRFTLEIPSYLPVMTYCNRALREEMYRAYATRA 244

Query: 244     SELSDD-GKFDNTANIDRTLENALQTAKLLGFKNYAELSLATKMADTPEQVINFLHDLAR 302
              SE + GK+DN+ ++ L ++ AKLLGF Y ELSLATKMA+ P+QVL+FL LA
Sbjct: 245     SEQGPNAGKWDNSKVMEEILTLRVELAKLLGFNTYTELSLATKMAENPQQVLDPLDHLAE 304

Query: 303     RAKPYAEKDLAEVKAFARESLGLADLPWDLGYAGEKLREAKYAFSETEVKKYFPVGVKL 362
              RAKP EK+L E+K + + G+ +L PWD+G+ EK ++ YA ++ E++ YFP +V+
Sbjct: 305     RAKPQGEKELQELKGYCEKEFGVTELPWDIGFYSEKQKQHLAINDIELRPYFPENRVI 364

```

Query: 363 NGLFAQIKKLYGIGFTE-KTVPVHKDVRYFEL-QQNETIGGVYMDLYAREGKRGAWM 420
 +GLF IK+++ I E K V WHKDVRF+L +N + G Y+DLYARE KRGAWM
 Sbjct: 365 SGLFELIKRIFNIRAVRKGVDTHKDVRFDDLIDENDQLRGSFYLDLYAREHKRGAWM 424

Query: 421 NDYKGRRRFSDGTLQLPTAYLVCNFTPPVGGKEARLSHDEIXXXXXXXXXXXXXXXXXQVD 480
 +D GR+R DG+++ P AYL CNF P+G K A +H+E+ Q+D
 Sbjct: 425 DDCIGRRKRLDGSJETPVAYLTCNFNAPIGNKPALFTHNEVTTLFHEFGHGIHMLTQID 484

Query: 481 ELGVSGINGVEWDAVELPSQFMENFVWEYNVLAQMSAHEETGVPLPKELFDKMLAAKNFQ 540
 V+GINGV WDAVELPSQFMEN+ WE LA +S H ETG PLPKE ++L AKNFQ
 Sbjct: 485 VSDVAGINGVPWDVELPSQFMENWCWEEALAFISGHYETGEPLPKEKLTQLLKAKNFQ 544

Query: 541 RGMFLVRQMEFALFDMMIYSEDDEGRLLKNWQQVLDVSRKEVAVVRPPEYNRFANSFGHIF 600
 MF++RQ+EF +FD ++ D + L SV+ +VAV++ ++ R +SF HIF
 Sbjct: 545 AAMFILRQLEFGIFDFRLHHTFDAEKTNIILDTLKS VKSQVAVIKGVWDWARAPHSFSHIF 604

Query: 601 XXXXXXXXXXXXWAEVLSADAYAFAFEESDDV-AATGKRFWQEILAVGGSRSAAESFKAFR 659
 WAEVLSADAY+ FEE TKG F EIL GGS E FK FR
 Sbjct: 605 AGGYAAGYYSYLWAEVLSADAYSRFEEGIFNPITGKSFLDEILTRGGSEEPMELFKFR 664

Query: 660 GREPSIDALLRHSGFDN 676
 GREP +DALLRH G N
 Sbjct: 665 GREPQLDALLRHKGIMN 681

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 493>:

g129.seq
 1 ATGCTTTCAC CTCCTCGGCG TAAACGGCG GCACATCAAT CAAGCCGTCT
 51 TTCATTGCG TCGGAAAAA ATGCGGCGTG TTGCCGTGAT CAAAATCAAT
 101 ATCGTGCAGC ATCCAGCCCA AATCGCGGTT TGCCTCGCTT TCCGATAACG
 151 CCGACGGCGG CAGCGGTTCA CCCTTATCCG CGCTTTCGCC ATTGCCCCTT
 201 TCAGGCTGCG GGCATAGGGG CGGAACAGGC GCGCGTCGAA TCCTGTTTCA
 251 TCCGGACAAA CGCGTTGGCA GTCGGAAAAA CCGGCCGGCC GTGTCAAATA
 301 ATGCGTTACT TTGGCCGGGT CTTGTCTTTT GTAAGCGGCG GTCTTTTTTT
 351 GCGCGCCATC CGCATCTGTT TGGGCGCATG GCAAACGGCG GCTGCCGTAC
 401 AATCAAAATG TTTGGCGATT TCATGCAGAC AGGCATCCGG ATGCCGCCCG
 451 ACATATCGAG CCGGTTTTTG CCTATCCGAT TTGGCGGCAT TAGGCCCGT
 501 AACTTGA

This corresponds to the amino acid sequence <SEQ ID 494; ORF 129.ng>:

g129.pep
 1 MLSPRRKTA AHQSSRLSFA CGKNAACCRD QNQYRAASSP NRGLPRFPIT
 51 PTAAAVHPYP RFRHLFPQAA GIGAEQAAVE SCFIRTNALA VGKSGRPCQI
 101 MRYFGRVLSF VSGGLFLRAI RICLGAWQTA AAVQSKCLAI SCRQASGCRP
 151 TYRAGFCLSD LAAPRPVT*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 495>:

m129.seq (partial)
 1 ..TATCTGCGCT TTTACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA
 51 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
 101 GAAAATTCGG CCGGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG
 151 TTCTTTGTAA GTGGTGGTCT TTTTTTGC GC GTTATCCCCA TCTGTTTGAG
 201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTTG GCGATTTCAT
 251 GCAGATAGGC ATCCGGGTGT TGCCCAACAT ATTGAGCCCG TTTTTCCTA
 301 TCCGATTGA CGGCATTAG ACCGGTAACT TGA

This corresponds to the amino acid sequence <SEQ ID 496; ORF 129>:

m129.pep (partial)
 1 ..YLRFHYPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGRLC QIMRYFGRVL
 51 FFVSGGLFLR VIPICLSAXQ MVAAVQSKCL AISCRXASGC CPTYXAGFCL
 101 SDLTAFRPVT *

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 129 shows 79.1% identity over a 110 aa overlap with a predicted ORF (ORF 129.ng) from *N. gonorrhoeae*:

m129/g129

```

                                10      20      30
m129.pep                      YLRFHYLPFQAAGIGTEQVAVKSCFIQINT
                                | | | : | | | | | | | | : | | : | | | : | :
g129      RDQNQYRAASSPNRGLPRFPITPTAAAVHPYPRFRHLPFQAAGIGAEQAAVESCFIRTNA
                                30      40      50      60      70      80

                                40      50      60      70      80      90
m129.pep      LVVGKFGRLCQIMRYFGRVLFVSGGLFLRVIPICLSAXQMVAAVQSKCLAISCRXASGC
                | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g129      LAVGKSGRPCQIMRYFGRVLSFVSGGLFLRAIRICLGAWQTAAAVQSKCLAISCRQASGC
                90      100     110     120     130     140

                                100     110
m129.pep      CPTYXAGFCLSDLTAFRPVTX
                | | | | | | | | : | | | | |
g129      RPTYRAGFCLSDLAAPRPVTX
                150     160

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 497>:

```

a129.seq (partial)
1  TATCTGCGCT TTCACTATTT GCCCTTTCAG GCTGCGGGCA TAGGGACGGA
51 ACAGGTAGCG GTCAAATCCT GTTTCATCCA AATAAACACG TTGGTAGTCG
101 GAAAATTCGG CCAGCTGTGT CAAATAATGC GTTACTTTGG CCGGGTCTTG
151 TTCTTTGTAA GTGGTGGTCT TTTTGTGCGC GTTATCCCCA TCTGTTTGAG
201 TGCATAGCAA ATGGTGGCTG CCGTACAATC AAAATGTTG GCGATTTTCAT
251 GCAGATAGGC ATCCTGGTGT TGCCCAACAT ATTGAGCCGG TTTTGCCTA
301 TCCGATTGA CGGCATTAG ACCGTAAC TGA

```

This corresponds to the amino acid sequence <SEQ ID 498; ORF 129.a>:

```

a129.pep (partial)
1  YLRFHYLPFQ AAGIGTEQVA VKSCFIQINT LVVGKFGQLC QIMRYFGRVL
51 FFVSGGLFLR VIPICLSA*Q MVAAVQSKCL AISCR*ASWC CPTY*AGFCL
101 SDLTAFRPVT *

```

m129/a129 98.2% identity in 110 aa overlap

```

                                10      20      30      40      50      60
m129.pep      YLRFHYLPFQAAGIGTEQVAVKSCFIQINTLVVGKFGRLCQIMRYFGRVLFVSGGLFLR
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a129      YLRFHYLPFQAAGIGTEQVAVKSCFIQINTLVVGKFGQLCQIMRYFGRVLFVSGGLFLR
                10      20      30      40      50      60

                                70      80      90      100     110
m129.pep      VIPICLSAXQMVAAVQSKCLAISCRXASGCCPTYXAGFCLSDLTAFRPVTX
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
a129      VIPICLSAXQMVAAVQSKCLAISCRXASWCCPTYXAGFCLSDLTAFRPVTX
                70      80      90      100     110

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 499>:

```

g130.seq
1  ATGAAACAAC TCCGCGACAA CAAAGCCCAA GGCTCTGCAC TGTTTACCTT
51 TGTGAGCGGT ATCGTTATTG TTATTGCAGT CCTTTATTTC CTGATTAAAGC
101 TGGCGGGCAG TGGATCGTTC GCGATGTCG ATGCCACTAC GGAAGCGGCA
151 ACGCAGACCC GCATCCAGCC TGTCGGACAA TTGACGATGG GTGACGGCAT
201 CCCCGTCGGC GAACGCCAAG GCGAACAGAT TTTCGGCAAA ATCTGTATCC

```

375

```

251 AATGCCACGC GCGGCACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGACT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCAC
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGCAG
401 ACCTGACCGA TCAGGAAGTC AAACGGGCGA TTACCTACAT GGCGAATAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGACGCT CCGGCAGAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651 CGGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAAGA AACCTTGCAC AAACATGCCC TTGAAGGCTT TAACGCGATG
751 CCGGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801 TGTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 500; ORF 130.ng>:

```

g130.pep
  1 MKQLRDNKAQ GSALFTLVSG IVIVIAVLYF LIKLAGSGSF GDVDATTEAA
 51 TQTRIQPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101 NGDWAPRIAQ GFDTLFQHAL NGFNAMPAKG GAADLTQDEL KRAITYMANK
151 SGGSFNPDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAPAVGVVDG
201 KKVFEATCQV CHGGSIPGIP GIGKKDDWAP RIKKGKETLH KHALEGFNAM
251 PAKGGNAGLS DDEVKAAVDY MANQSGAKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 501>:

```

m130.seq (partial)
  1 ..GGCGAACAGA TTTTCGGCAA AATCTGTATC CAATGCCACG CCGCGGACAG
 51 CAATGTGCCG AACGCTCCGA AACTGGAACA CAACGGCGAT TrGGCACCGC
101 GTATCGGCAA GGCTTCGATA CCTTGTTCAC ACACGCGCTG AACGGCTTTA
151 ACGCCATGCC TGCAAAAGGC GGTGCGGCAG ACCTGACCGA TCAGGAAGTT
201 AAACGGGCGA TTACTTACAT GGCGAACAAA AGCGGCGGTT CTTTCCCGAA
251 TCCTGATGAG GCTGCGCCTG CCGACAATGC CGCTTCAGGA ACAGCTTCTG
301 CTCCTGCCGA TAGTGACGCT CCGGCAGAAG CGAAGGCAGA AGACAAGGGT
351 CCGGCACCCC TGCGGTCGGC GTTGACGGTA AAAAAGTCTT CGAAGCAACC
401 TGTCAGGTGT GCCACGGCGG TTCGATTCCC GGTATTCCCG GCATAGGCAA
451 AAAAGACGAT TGGGCACCGC GTATCAAAAA AGGCAAAGAA ACCTGCACA
501 AACACGCCCT TGAAGGCTTT AACGCGATGC CTGCCAAArG CGgCAATGCA
551 GGTTTGAGCG ATGACGAAGt CAAAGCGGCT GTTGACTATA TGGCAAACCA
601 ATCCGGTGCA AAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 502; ORF 130>:

```

m130.pep (partial)
  1 ..GEQIFGKICI QCHAADSNVP NAPKLEHNGD XAPRIQGFDL LFQHALNGFN
 51 AMPAKGGAAD LTDQELKRAI TYMANKSGGS FPNPDEAAPA DNAASGTASA
101 PADSAPAEA KAEDKGAAPA VGVDGKKVFE ATCQVCHGGS IPGIPGIGKK
151 DDWAPRIKKG KETLHKHALE GFNAMPKXG NAGLSDEVK AAVDYMANQS
201 GAKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 130 shows 98.1% identity over a 206 aa overlap with a predicted ORF (ORF 130.ng) from *N. gonorrhoeae*:

m130/g130

```

                                     10      20      30
m130.pep                               GEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                     |||
g130      DATTEAATQTRIQPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
                                     50      60      70      80      90     100

                                     40      50      60      70      80      89
m130.pep      XAPRI-QGFDTLFQHALNGFNAMPAKGGAADLTQDELKRAITYMANKSGGSFPNPDEAAP
               |||
g130      WAPRIAQGFDTLFQHALNGFNAMPAKGGAADLTQDELKRAITYMANKSGGSFPNPDEAAP
               110     120     130     140     150     160

```

376

```

          90      100      110      120      130      140
m130.pep  ADNAASGTASAPADSAAPAEAKAEDKGAA-PAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          |||
g130       ADNAASGTASAPADSAAPAEAKAEDKGAAAPAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          170      180      190      200      210      220

          150      160      170      180      190      200
m130.pep  KKDDWAPRIKKGKETLHKHALEGFNAMPKXGNAGLSDDDEVKAAVDYMANQSGAKFX
          |||
g130       KKDDWAPRIKKGKETLHKHALEGFNAMPKGGNAGLSDDDEVKAAVDYMANQSGAKFX
          230      240      250      260      270      280

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 503>:

```

a130.seq
1  ATGAAACAAC TCCGCGACAA CAAAGCCCAA GGCTCTGCAC TGTTTACCCT
51  TGTGAGCGGT ATCGTTATTG TTATTGCAGT CCTTTATTTC CTGATTAAGC
101 TGGCGGGCAG CGGCTCGTTC GCGGATGTCG ATGCCACTAC GGAAGCAGCA
151 ACGCAGACCC GTATCCAGCC TGTCGGACAA TTGACGATGG GCGACGGCAT
201 CCCCCTCGGC GAACGCCAAG GCGAACAGAT TTTCGGCAAA ATCTGTATCC
251 AATGCCACGC GCGGCACAGC AATGTGCCGA ACGCTCCGAA ACTGGAACAC
301 AACGGCGATT GGGCGCCGCG TATCGCGCAA GGCTTCGATA CCTTGTTCCA
351 ACACGCGCTG AACGGCTTTA ACGCCATGCC TGCCAAAGGC GGTGCGGTAG
401 ACCTGACCGA TCAGGAAGTC AAACGGGCGA TTAATTACAT GGCGAACAAA
451 AGCGGCGGTT CTTTCCCGAA TCCTGATGAG GCTGCGCCTG CCGACAATGC
501 CGCTTCAGGA ACAGCTTCTG CTCCTGCCGA TAGTGCAGCT CCGGCAGAAG
551 CGAAGGCAGA AGACAAGGGT GCGGCAGCCC CTGCGGTCGG CGTTGACGGT
601 AAAAAAGTCT TCGAAGCAAC CTGTCAGGTG TGCCACGGCG GTTCGATTCC
651 CGGTATTCCC GGCATAGGCA AAAAAGACGA TTGGGCACCG CGTATCAAAA
701 AAGGCAAAGA AACCTTGCAC AAACACGCCC TTGAAGGCTT TAACGCGATG
751 CCTGCCAAAG GCGGCAATGC AGGTTTGAGC GATGACGAAG TCAAAGCGGC
801 TGTTGACTAT ATGGCAAACC AATCCGGTGC AAAATTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 504; ORF 130.a>:

```

a130.pep
1  MKQLRDNKAQ GSALFTLVSG IVIVIAVLYF LIKLAGSGSF GDVDATTEAA
51  TQTRIQPVGQ LTMGDGIPVG ERQGEQIFGK ICIQCHAADS NVPNAPKLEH
101 NGDWAPRIAQ GFDTLFQHAL NGFNAMPAKG GAVDLTDQEL KRAITYMANK
151 SSGSFPNPDE AAPADNAASG TASAPADSAA PAEAKAEDKG AAAPAVGVVDG
201 KKVFEATCQV CHGGSIPGIP GIGKKDDWAP RIKKGKETLH KHALEGFNAM
251 PAKGGNAGLS DDEVKAAVDY MANQSGAKF*

```

m130/a130 97.6% identity in 206 aa overlap

```

          10      20      30
m130.pep  GEQIFGKICIQCHAADSNVPNAPKLEHNGD
          |||
a130       DATTEAATQTRIQPVGQLTMGDGIPVGERQGEQIFGKICIQCHAADSNVPNAPKLEHNGD
          50      60      70      80      90      100

          40      50      60      70      80      89
m130.pep  XAPRI-QGFDTLFQHALNGFNAMPAKGGAADLTDQELKRAITYMANKSGGSFPNPDEAAP
          |||
a130       WAPRIAQGFDTLFQHALNGFNAMPAKGGAVDLTDQELKRAITYMANKSGGSFPNPDEAAP
          110      120      130      140      150      160

          90      100      110      120      130      140
m130.pep  ADNAASGTASAPADSAAPAEAKAEDKGAA-PAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          |||
a130       ADNAASGTASAPADSAAPAEAKAEDKGAAAPAVGVDGKKVFEATCQVCHGGSIPGIPGIG
          170      180      190      200      210      220

          150      160      170      180      190      200
m130.pep  KKDDWAPRIKKGKETLHKHALEGFNAMPKXGNAGLSDDDEVKAAVDYMANQSGAKFX

```

377

```

|||||
a130      KKDDWAPRIKKKGT LHKH ALEGFNAMP AKGGNAGLS DDEVKAAVDYMANQSGAKFX
           230       240       250       260       270       280

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 505>:

```

g132.seq
1  ATGGAAGCCT TCAAAACCCCT AATTTGGATT ATTAATATTA TTTCCGCTTT
51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACCTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151 GCCGGCAACG CCAACTTcct CagccGCTCG AccGccGTTG CAGCAACAtt
201 tttcttTGca acctgcAtgg gctatggTgt atattcacac CCACACGACA
251 AAACACGGTT TGGACTtcag caacataCGA CAGACTCAGC AagcACCCAA
301 ACCcgtAAGC AATACCGAAC CTTCTGCCCC TGTTCCTCAG CAGCAGAAAT
351 AACagtTTTT CAAATgccga caTGgtga

```

This corresponds to the amino acid sequence <SEQ ID 506; ORF 132.ng>:

```

g132.pep
1  MEAFKTLIWI INIISALAVI VLVL LQHKGK ADAGATFGSG SGSAQGVFGS
51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QHTTDSASTQ
101 TRKQYRTFCP CSSAAEITVF QMPTW*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 507>:

```

m132.seq (partial)
1  ATGGAACCCCT TCAAAACCTT AATTTGGATT GTTAATTTAA TTTCCGCTTT
51  GGCCGTCTTC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACTTT CGGA...

```

This corresponds to the amino acid sequence <SEQ ID 508; ORF 132>:

```

m132.pep (partial)
1  MEPFKTLIWI VNLISALAVF VLVL LQHKGK ADAGATFG...

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 132 shows 89.5% identity over a 38 aa overlap with a predicted ORF (ORF 132.ng) from *N. gonorrhoeae*:

```

m132/g132
           10       20       30
m132.pep  MEPFKTLIWI VNLISALAVF VLVL LQHKGK ADAGATFG
           || ||||| : ||||| : ||||| ||||| |||||
g132      MEAFKTLIWI INIISALAVI VLVL LQHKGK ADAGATFGSGSGSAQGVFGSAGNANFLSRS
           10       20       30       40       50       60

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 509>:

```

a132.seq
1  ATGGAAGCCT TCAAAACCCCT AATTTGGATT GTTAATATTA TTTCCGCTTT
51  GGCCGTCATC GTGTTAGTAT TGCTCCAACA CGGCAAAGGC GCGGATGCCG
101 GCGCGACTTT CGGATCGGGA AGCGGCAGCG CGCAAGGCGT ATTCGGCTCT
151 GCCGGCAACG CTAACCTCCT CAGCCGCTCG ACCGCCGTTG CAGCAACATT
201 TTTCTTTGCA ACCTGCATGg GCTATGGTGT ATATTACAC CCACACGACA
251 AAACACGGTT TGGACTTCAG CAACGTACAA CAAACTCAGC AAGCACCCAA
301 ACCCGTAAGC AATACCGAAC CTTCTGCCCC TGTTCCTCAG CAGCAGAAAT
351 AACAGTTTTT CAAATGCCGA CATGGTGA

```

This corresponds to the amino acid sequence <SEQ ID 510; ORF 132.a>:

```

a132.pep
1  MEAFKTLIWI VNIISALAVI VLVL LQHKGK ADAGATFGSG SGSAQGVFGS
51  AGNANFLSRS TAVAATFFFA TCMGYGVYSH PHDKTRFGLQ QRTTNSASTQ
101 TRKQYRTFCP CSSAAEITVF QMPTW*

```

m132/a132 92.1% identity in 38 aa overlap

378

	10	20	30	
m132.pep	MEPFKTLIWIIVNLISALAVFVLVLLQHGKGADAGATFG			
	: :			
a132	MEAFKTLIWIIVNIISALAVIVLVLLQHGKGADAGATFGSGSGSAQGVFGSAGNANFLSRS			
	10	20	30	40 50 60

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 511>:

g134.seq

```

1  ATGTCCCAAG AAATCCTCGA CCAAGTGC GC CGCCGCCGCA CGTTTGCCAT
51  CATCTCCAC CCCGATGCGG GTAAAACAC GCTGACCGAA AAACCTGCTGC
101 TGTTCGCGG CGCGATTCAA AGCGCAGGCA CGGTGAAAGG TAAGAAAACC
151 GGCAAATTCG CCACCTCCGA CTGGATGGAC ATCGAGAAGC AGCGCGGCAT
201 TTCCGTGGCA TCAAGCGTGA TGCAGTTCGA CTACAAAGAC CACACCGTCA
251 ACCTCTTGGA CACGCCGCGG CACCAAGACT TCTCCGAAGA CACCTACCGC
301 GTTTTAACCG CAGTGGACAG CGCCTTGATG GTCATCGACG CGGCAAAAGG
351 CGTGGGAAGCG CAAACCATCA AACTCTTGAA CGTCTGCCGC CTGCGCGATA
401 CGCCGATTGT TACCTTCATG AACAAATACG ACCGCGAAGT GCGCGATTCT
451 TTGGAACCTT TGGACGAAGT GGAAGACATC CTGCAAATCC GCTGCGCGCC
501 CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAGGGC GTGTACCACA
551 TCCTGAACGA CGAAATCTAT CTCTTGAAAG CGGGCGGCGA ACGCCTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATAAAC AATCCCGAAT TGGAAACAACG
651 CTTTCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGGCTTCCAA CGAATTTAAT CTCGacgaAT TTCTCGccgG CGAACTCACG
751 CCAGTGTCTT TCGGCTCTGC GATTAACAAC TTCGGCATTC AGGAAATCCT
801 CAATTCATTG ATTGACTGGG CACCCGCACC GAAACCGCGC GACGCGACCA
851 TGGCATGGT CGGGCCGGAC GAGCCGAAAT TTCCCGATT TATCTTTAAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATCG CCTTCTTGCG
951 CGTCTGCTCC GGTAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTAG TAACCTTCAT GTCGCACGAC
1051 CGCGAACTGG CGGAAGAAGC CTACGCCGGC GACATCATCG GCATCCCGAA
1101 CCACGGCAAC ATCCAATCG GCGACAGCTT CTCCGAAGGC GAACAACATG
1151 CGTTTACCGG CATCCCATTC TTCGCGCCCG AACTGTTCCG CAGCGTCCGC
1201 ATCAAAACC CGTGAAAAT CAAACAACTG CAAAAGGTT TGCAACAAC
1251 CGGCGAAGAA GGTGCGGTTT AAGTATTCAA ACCGATGAGC GGCGCGGATT
1301 TGATTTTGGG TCGGGTCGGC GTGTTGCAGT TTGAAGTCGT AACCTCACGC
1351 CTCGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAGCG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAACTG GCGGAATTG
1451 AAAAAGCCAA CGCAGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501 TACCTCGCCC CCAACCGCGT GAATTGCGG TTGACGCAAG AACGCTGCCC
1551 GGACATCGTG TTCCACGAAA CGCGCGAACA TTCGGTCAAA CTCTAA

```

This corresponds to the amino acid sequence <SEQ ID 512; ORF 134.ng>:

g134.pep

```

1  MSQEILDQVR RRRTFAIISH PDAGKTLTE KLLLFSGAIQ SAGTVKGGKT
51  GKFATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAGVEA QTIKLLNVCR LRDTPIVTFM NKYDREVRDS
151 LELLDEVEDI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGIN NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATMRMVGPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELAEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVTSR
451 LANEYGVAV FDSASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 513>:

m134.seq

```

1  ATGTCCCAAG AAATCCTCGA CCAAGTGC GC CGCCGCCGCA CGTTTGCCAT
51  CATCTCCAC CCGACGCGAG GTAAAACAC GTTGAAGTAA AAACCTTTCG
101 TGTTCGCGG CGCGATTCAA AGCGCGGGTA CGGTAAAAGG CAAGAAAACC
151 GGCAAATTCG CCACTCCGA CTGGATGGAA ATCGAGAAGC AGCGCGGCAT

```


379

```

201 TTCCGTGGCA TCAAGTGTGA TGCAGTTCGA TTACAAAGAC CACACCGTCA
251 ACCTCTTGGA CACGCCGGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301 GTTTTAACCG CCGTGGACAG CGCATTAATG GTCATCGACG CGGCAAAAGG
351 CGTGGGAAGCG CAAACCATCA AGCTCTTAAA CGTCTGCCGC CTGCGCGATA
401 CACCGATTGT TACGTTTATG AACAAATACG ACCGCGAAGT GCGCGATTCC
451 CTGGAACTTT TGGACGAAGT GGAAACATT TTAATAATCC GCTGCGCGCC
501 CGTTACCTGG CCGATCGGTA TGGGCAAAAA CTTCAAGGGC GTGTACCACA
551 TCCTGAACGA TGAAATTTAT CTCTTGAAAG CTGGCGGCGA ACGCCTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCTGAAT TGGAACAACG
651 CTTTCCGTTG GAAATCCAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGGCTTCCAA CGAGTTTAAT CTCGACGAAT TCCTCGCCGG CGAACTCAGC
751 CCCGTATTCT TCGGCTCTGC GATTAACAAC TTCGGTATTC AGGAAATCCT
801 CAATTCATTG ATTGACTGGG CGCCCGCGCC GAAACCGCGC GACGCGACCG
851 TACGTATGGT CGAGCCGGAC GAGCCGAAGT TTTCCGGATT TATCTTCAAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CCTTCTTGCG
951 CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAGATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TTACCTTCAT GTCGCACGAC
1051 CGCGAGCTGG TTGAAGAAGC CTACGCCGGC GACATTATCG GCATCCCGAA
1101 CCACGGCAAC ATCCAATCG GCGACAGCTT CTCGAAGGC GAACAACCTG
1151 CGTTCACCGG CATCCCATTC TTCGCACCCG AACTGTTCCG CAGCGTACGC
1201 ATCAAAAACC CGCTGAAAT CAAACAACCTG CAAAAAGGCT TGCAACAGCT
1251 CGGCGAAGAA GCGCGGTGTC AGGTGTTCAA ACCGATGAGC GCGCGGATT
1301 TGATTTTGGG CGCGGTGCGC GTGTGTCAGT TTGAAGTCGT TACCTCGCGC
1351 CTCGCCAACG AATACGGCGT AGAAGCCGTG TTCGACAGCG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAATG GCTGAATTTG
1451 AAAAAGCCAA CGCGGGCAAC CTCGCCATCG ACGCAGGCGG CAACCTCGCC
1501 TACCTCGCCC CCAACCGCGT GAATTTGGGA CTCACGCAAG AACGTTGGCC
1551 GGACATCGTG TTCCACGAAA CACGCGAACA TTCGGTCAAA CTGTAA

```

This corresponds to the amino acid sequence <SEQ ID 514; ORF 134>:

```

m134.pep
1 MSQEILDQVR RRRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
51 GKFATSDWME IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAGVEA QTIKLLNVCR LRDTPIVTFM NKYDREV RDS
151 LELLDEVENI LKIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDI IKGID NPELEQRFPL EIQQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IDWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFM SHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLAFTGIFP FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGV EAV FDSASIWSAR WVSCDDKKKL AFEKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 134 shows 98.7% identity over a 531 aa overlap with a predicted ORF (ORF 134.ng)

from *N. gonorrhoeae*:

m134/g134

```

          10      20      30      40      50      60
m134.pep MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
          |||
g134      MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD
          10      20      30      40      50      60

          70      80      90      100     110     120
m134.pep IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA
          |||
g134      IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA
          70      80      90      100     110     120

          130     140     150     160     170     180

```

380

m134 . pep	QTIKLLNVCLRDTPIVTFMNKYDREVRDSLELLDEVENILKIRCAPVTWPIGMGNFKG
g134	QTIKLLNVCLRDTPIVTFMNKYDREVRDSLELLDEVEDILQIRCAPVTWPIGMGNFKG
	130 140 150 160 170 180
m134 . pep	VYHILNDEIYLF EAGGERLPHEFDI IKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
g134	VYHILNDEIYLF EAGGERLPHEFDI IKGINNPELEQRFPLEIQQLRDEIELVQAASNEFN
	190 200 210 220 230 240
m134 . pep	LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATVRMVEPDEPKFSGFIFK
g134	LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATMRMVGPDPEPKFSGFIFK
	250 260 270 280 290 300
m134 . pep	IQANMDPKHRDRIAFLRVC SGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
g134	IQANMDPKHRDRIAFLRVC SGKFERGMKMKHLRINREIAASSVVTFMSHDRELAEEAYAG
	310 320 330 340 350 360
m134 . pep	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIQQLQKGLQQLGEE
g134	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIQQLQKGLQQLGEE
	370 380 390 400 410 420
m134 . pep	GAVQVFKPMMSGADLILGAVGV LQFEVVTSRLANEYGV EAVFDSASISARWVSCDDKKKL
g134	GAVQVFKPMMSGADLILGAVGV LQFEVVTSRLANEYGV EAVFDSASISARWVSCDDKKKL
	430 440 450 460 470 480
m134 . pep	AEFEKANAGNLAIDAGGNLAY LAPNRVNLGLTQERWPDIVFHETREHSVKLX
g134	AEFEKANAGNLAIDAGGNLAY LAPNRVNLGLTQERWPDIVFHETREHSVKLX
	490 500 510 520 530

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 515>:

```

a134.seq
1  ATGTCCCAAG AAATCCTCGA CCAAGTGCGC CGCCGCCGCA CGTTTGCCAT
51 CATCTCCAC CCTGACGCAG GTAAAACCAC GTTGACTGAA AAACCTTTGC
101 TGTTTTCAGG TGCATTCAA AGCGCGGGTA CGGTAAAAGG CAAGAAAACC
151 GGCAAATTCG CCACCTCCGA CTGGATGGAC ATCGAGAAGC AGCGCGGCAT
201 TTCCGTGGCA TCAAGCGTGA TGCAGTTCGA CTATAAGAC CACACCGTCA
251 ACCTTTTGGA CAGCCCGGGA CACCAAGACT TCTCCGAAGA CACCTACCGC
301 GTTTTGACCG CCGTCGATAG TGCCTTGATG GTCATCGACG CGGCAAAAGG
351 CGTGGGAAGCG CAAACCATCA AACTCTTGAA CGTCTGCCGC CTGCGCAATA
401 CGCCGATTGT TACGTTTCATG AACAAATACG ACCGCGAAGT GCGCGATTCC
451 CTGGAATTGC TGGACGAAGT GGAAAACATC CTGCAAATCC GCTGCGCGCC
501 CGTAACCTGG CCGATCGGCA TGGGCAAAAA CTTCAAAGGC GTGTACCACA
551 TCCTGAACGA CGAAATCTAT CTCTTGAAAG CGGGCGGCGA ACGCTTGCCG
601 CACGAGTTCG ACATCATCAA AGGCATCGAT AATCCCGAAT TGGAACAACG
651 CTTTCCGTTA GAAATACAGC AGTTGCGCGA CGAAATCGAA TTGGTGCAGG
701 CGGCTTCCAA CGAGTTCAAT CTCGACGAAT TCCTCGCCGG CGAACTCACG
751 CCCGTATTCT TCGGCTCTGC GATTAACAAC TTCGTATTG AGGAAATCCT
801 CAATTCATTG ATTGAATGGG CGCCCGCGCC GAAACCACGC GATGCGACCG
851 TCGGTATGGT CGAGCCGGAC GAGCCGAAGT TTTCCGGATT TATCTTCAA
901 ATCCAAGCCA ATATGGACCC GAAACACCGC GACCGTATTG CCTTCTTGCG

```

381

```

951 CGTCTGCTCC GGCAAATTCG AGCGCGGCAT GAAAATGAAA CACCTGCGTA
1001 TCAACCGCGA AATCGCCGCC TCCAGCGTGG TAACCTTCAT GTCCACGAC
1051 CGCAGAGCTG TTGAAGAAGC CTACGCCGGC GACATTATCG GTATCCCAA
1101 CCACGGCAAC ATCCAAATCG GCGACAGCTT CTCCGAAGGC GAACAACTGA
1151 CGTTTACCGG CATCCCATTG TTCGCGCCCG AACTGTTCCG CAGCGTTCGC
1201 ATCAAAAACC CGCTGAAAAT CAAGCAACTG CAAAAAGGTT TGCAACAGCT
1251 TGGCGAAGAA GGTGCGGTGC AGGTGTTCAA ACCAATGAGC GGCGCGGATT
1301 TGATTTTGGG CGCGGTCGGC GTGTTGCAGT TTGAAGTCGT TACCTCGCGC
1351 CTTGCCAACG AATACGGCGT GGAAGCCGTG TTCGACAACG CATCCATCTG
1401 GTCGGCGCGC TGGGTATCGT GCGACGACAA GAAAAAAGT GCGGAATTTG
1451 AAAAAAGCAA CGCGGGCAAC CTCGCCATCG ACGCGGCGCG CAACCTCGCC
1501 TACCTCGCCC CTAACCGCGT GAATCTGGGA CTCACGCAAG AACGCTGGCC
1551 GGACATCGTG TTCCACGAAA CGCGCGAGCA TTCGGTCAAA CTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 516; ORF 134.a>:

```

a134.pep
1 MSQEILDQVR RRRTFAIISH PDAGKTTLTE KLLLFSGAIQ SAGTVKGKKT
51 GKFATSDWMD IEKQRGISVA SSVMQFDYKD HTVNLLDTPG HQDFSEDYR
101 VLTAVDSALM VIDAAGVEA QTIKLLNVCRLRNTPIVTFM NKYDREVRDS
151 LELLDEVENI LQIRCAPVTW PIGMGKNFKG VYHILNDEIY LFEAGGERLP
201 HEFDIIKGID NPELEQRFPL EIQLRDEIE LVQAASNEFN LDEFLAGELT
251 PVFFGSAINN FGIQEILNSL IEWAPAPKPR DATVRMVEPD EPKFSGFIFK
301 IQANMDPKHR DRIAFLRVCS GKFERGMKMK HLRINREIAA SSVVTFMSHD
351 RELVEEAYAG DIIGIPNHGN IQIGDSFSEG EQLTFTGIPF FAPELFRSVR
401 IKNPLKIKQL QKGLQQLGEE GAVQVFKPMS GADLILGAVG VLQFEVVTSR
451 LANEYGEAV FDNASIWSAR WVSCDDKKKL AEFKANAGN LAIDAGGNLA
501 YLAPNRVNLG LTQERWPDIV FHETREHSVK L*

```

m134/a134 98.9% identity in 531 aa overlap

```

m134.pep      10      20      30      40      50      60
MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWME
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a134          10      20      30      40      50      60
MSQEILDQVRRRTFAIISHPDAGKTTLTEKLLLFSGAIQSAGTVKGKKTGKFATSDWMD

m134.pep      70      80      90      100     110     120
IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a134          70      80      90      100     110     120
IEKQRGISVASSVMQFDYKDHTVNLLDTPGHQDFSEDYRVLTAVDSALMVIDAAGVEA

m134.pep      130     140     150     160     170     180
QTIKLLNVCRLRDTPIVTFMKNYDREVRDSLELLDEVENILKIRCAPVTWPIMGKNFKG
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a134          130     140     150     160     170     180
QTIKLLNVCRLRNTPIVTFMKNYDREVRDSLELLDEVENILQIRCAPVTWPIMGKNFKG

m134.pep      190     200     210     220     230     240
VYHILNDEIYLF EAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a134          190     200     210     220     230     240
VYHILNDEIYLF EAGGERLPHEFDIIKGIDNPELEQRFPLEIQQLRDEIELVQAASNEFN

m134.pep      250     260     270     280     290     300
LDEFLAGELTPVFFGSAINNFGIQEILNSLIDWAPAPKPRDATVRMVEPD EPKFSGFIFK
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a134          250     260     270     280     290     300
LDEFLAGELTPVFFGSAINNFGIQEILNSLIEWAPAPKPRDATVRMVEPD EPKFSGFIFK

m134.pep      310     320     330     340     350     360
IQANMDPKHRDRIAF LRVC SGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
a134          310     320     330     340     350     360
IQANMDPKHRDRIAF LRVC SGKFERGMKMKHLRINREIAASSVVTFMSHDRELVEEAYAG

```

382

	370	380	390	400	410	420
m134.pep	DIIGIPNHGNIQIGDSFSEGEQLAFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE					
a134	DIIGIPNHGNIQIGDSFSEGEQLTFTGIPFFAPELFRSVRIKNPLKIKQLQKGLQQLGEE					
	370	380	390	400	410	420
	430	440	450	460	470	480
m134.pep	GAVQVFKPMMSGADLILGAVGVLFQEVVTSRLANEYGVAVFDSASIWSARWVSCDDKKKL					
a134	GAVQVFKPMMSGADLILGAVGVLFQEVVTSRLANEYGVAVFDNASIWSARWVSCDDKKKL					
	430	440	450	460	470	480
	490	500	510	520	530	
m134.pep	AEFEKANAGNLAI DAGGNLAYLAPNRVNLGLTQERWPDIVFHETREHSVKLX					
a134	AEFEKANAGNLAI DAGGNLAYLAPNRVNLGLTQERWPDIVFHETREHSVKLX					
	490	500	510	520	530	

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 517>:

g135.seq

```

1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCG
51  TTCGGAcgGC AGCCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCTG CAGGGTTCGG GCGCTGGGT TTCAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GTGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCGATTCCC ATCATCAATG
401 AAAACGACAC GGTTCGGTT GAGGAGTTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCAGT
551 CGGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCCGG AAACGGCACG GCGGTATGTC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCACTCAA ACCCGATTCA TTGGCCGAAG CCGCCGAACA TCAGGCGGAC
751 GGCTCGTTTT TCGTcccCcg tgCCAAAGGT TTGCGGACAC AGAAGCAATG
801 GctggCGTTC TATTCCgaaa gcggGGgcag cggtTAtgtg gacgaaagtg
851 cggaaacacgc tTgtccgaa caagggaaaag cctgCTGA

```

This corresponds to the amino acid sequence <SEQ ID 518; ORF 135.ng>:

g135.pep

```

1  MKYKRIVFKV GTSSITRSDG SLSRGKIQTI TRQLAALHHA GHELVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAIP IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHINHEIIEM
201 AGGSGSANGT GGMLTKIKAA TIAESGVPV YICSSLKPDS LAEAAEHQAD
251 GSFFVPRAGK LRTQKQWLAF YSESGSVVY DESAEHALSE QGKAC*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 519>:

m135.seq

```

1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51  TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCTGCCAGC
101 TTGCCGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTGCGG CAGGGTTCGG TCGCTGGGT TTCAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACCTG TCTTCAGACG GCATCGTGTC CGCGCAAATC
301 CTGCTCAGCC GCGCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
401 AAAACGATAC GGTTCGGTT GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGCTGCT

```

```

501 GACCGACATA GACGGTCTTT ACACGGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 GCGGGCGGCT CGGGTTCGGC AACCGGCACG GCGGGTATGC TGACCAAAAT
651 CAAAGCGGCA ACCATCGCCG CCGAATCCGG CGTACCGGTG TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CTGCCGAACA TCAGGCGGAC
751 GGCTCGTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATTCCGAAA GCCGGGGCAG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTTGTCCGAA CAGGGGAAAA GCCTGCTGAT GTCGGGCATT
901 GCCGAATCG AAGGGCATTT TTCCCGTATG GACACCGTAA CCGTGACAG
951 CAAGGCAACC AAACAGCCCC TGGGCAAAGG GCGCGTCCTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATCGCGTA AGGCGAAAGG CGTGTTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTTAG

```

This corresponds to the amino acid sequence <SEQ ID 520; ORF 135>:

```

m135.pep
  1 MKYKRIVFKV GTSSITHSDG SLRSGKIQTI TCQLAALHHA GHELVLVSSG
 51 AVAAGFGALG FKKRPVKIAD KQASAAVGQG LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRAPV IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGPNPS NPDAVRLDKI EHNHEIEM
201 AGGSGSANGT GGMLTKIKAA TIAAESGVFV YICSSLKPDA LAEAAEHQAD
251 GSFFVPRAGK LRTQKQWLAF YSESRSVYV DEGAEHALSE QGKSLLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAAE DLL KSRKAKGVFI
351 HRDDWISITP EIRLLLTEF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 135 shows 97.6% identity over a 294 aa overlap with a predicted ORF (ORF 135.ng) from *N. gonorrhoeae*:

```

m135/g135
      10      20      30      40      50      60
m135.pep  MKYKRIVFKVGTSSITHSDGSLRSGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      MKYKRIVFKVGTSSITRSDGSLRSGKIQTITRQLAALHHAGHELVLVSSGAVAAGFGALG
          10      20      30      40      50      60

      70      80      90     100     110     120
m135.pep  FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQIILSRADFADKRRYQNAGGAL
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      FKKRPVKIADKQASAAVGQGLLMEEYTANLSSDGIVSAQIILSRADFADKRRYQNAGGAL
          70      80      90     100     110     120

      130     140     150     160     170     180
m135.pep  SVLLQRRAPVIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGPNPS
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      SVLLQRRAPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGPNPS
          130     140     150     160     170     180

      190     200     210     220     230     240
m135.pep  NPDAVRLDKIEHNHEIEMAGGSGSANGTGGMLTKIKAATIAESGVFVYICSSLKPDA
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      NPDAVRLDKIEHNHEIEMAGGSGSANGTGGMLTKIKAATIAESGVFVYICSSLKPD
          190     200     210     220     230     240

      250     260     270     280     290     300
m135.pep  LAEAAEHQADGSFFVPRAGLRTQKQWLAFYSESRSVYVDEGAEHALSEQGKSLLMSGI
          |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g135      LAEAAEHQADGSFFVPRAGLRTQKQWLAFYSESRSVYVDESAEHALSEQGKACX
          250     260     270     280     290

      310     320     330     340     350     360

```

m135.pep AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAEDLLKSRKAKGVFIHRDDWISITP

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 521>:

```
a135.seq
1  ATGAAATACA AAAGAATCGT ATTTAAAGTC GGCACATCTT CGATTACCCA
51  TTCGGACGGC AGTCTCTCGC GCGGCAAAAT CCAAACCATC ACCCGCCAGC
101 TTGCGGCATT GCATCATGCG GGACACGAGC TGGTCTTGGT GTCTTCCGGC
151 GCGGTTCGCG CAGGGTTCGG TGCGCTGGGT TTCAAAAAAC GTCCGGTCAA
201 AATCGCCGAC AAACAGGCTT CCGCCGCCGT CGGGCAGGGG CTGCTGATGG
251 AAGAATATAC GGCAAACTG TCTTCAGACG GCATCGTGTC CGCACAAATC
301 CTGCTCAGCC GCGCCGACTT TGCCGACAAA CGCCGCTACC AAAATGCCGG
351 CGGCGCACTT TCCGTGCTGC TGCAACGCCG CGCCGTCCCC ATCATCAATG
401 AAAACGATAC GGTTCGGTT GAGGAATTGA AAATCGGCGA CAACGACACA
451 TTGAGTGCGC AAGTGGCGGC GATGATACAG GCAGACCTCT TGGTGTCTGT
501 GACCGACATA GACGGTCTTT ACACCGGCAA CCCGAACAGC AATCCCGATG
551 CCGTACGGCT GGACAAAATC GAACACATCA ACCATGAAAT CATCGAAATG
601 CCGGGCGGCT CGGGTTCGGC AAACGGCACA GGCGGTATGC TGAATAAATG
651 CAAAGCGGCG ACGATTGCGA CCGAGTCCGG CGTACCGGTC TATATCTGTT
701 CCTCGCTCAA ACCCGATGCA CTTGCCGAAG CGGCAGATAA TCAGGCGGAC
751 GGTCGTTTTT TCGTCCCCCG TGCCAAAGGT TTGCGGACGC AGAAGCAATG
801 GCTGGCGTTC TATTCGAAA GCAGGGGCGG CGTTTATGTG GACGAAGGTG
851 CGGAACACGC TTTGTCCGAA CAGGGAAAAA GCCTGCTGAT GTCGGGCATT
901 GCCGGAATCG AAGGGCATT TTCCCGTATG GACACCGTAA CCGTGTACAG
951 CAAGGCAACC AAACAGCCTT TGGGCAAAGG GCGAGTCTTG TTCGGCTCTG
1001 CCGCCGCCGA AGACCTGCTC AAATTGCGTA AGGCGAAAGG CGTGTTCATC
1051 CATCGGGACG ACTGGATTTC CATCACGCC GAAATACGCC TGCTTCTGAC
1101 CGAATTTTAG
```

This corresponds to the amino acid sequence <SEQ ID 522; ORF 135.a>:

```
a135.pep
1  MKYKRIVFKV GTSSITHSDG SLSRGKIQTI TROLAALHHA GHLEVLVSSG
51  AVAAGFGALG FKKRPVKIAD KQASAAVGGQ LLMEEYTANL SSDGIVSAQI
101 LLSRADFADK RRYQNAGGAL SVLLQRRVAV IINENDTVSV EELKIGDNDT
151 LSAQVAAMIQ ADLLVLLTDI DGLYTGNPNS NPDAVRLDKI EHNHEIEM
201 AGGSGSANGT GMLTKIKAA TIATESGVPV YICSSLKPDA LAEAADNQAD
251 GSFFVPRAKG LRTQKQWLAF YSESRRGVYV DEGAEHALSE QGKSLLMSGI
301 AGIEGHFSRM DTVTVYSKAT KQPLGKGRVL FGSAAEDLL KLRKAKGVFI
351 HRDDWISITP EIRLLLTEF*
```

m135/a135 98.4% identity in 369 aa overlap

	10	20	30	40	50	60
m135.pep	MKYKRIVFKVGTSSITHSDGSLSRGKIQTITCQLAALHHAGHELVLVSSGAVAAGFGALG					
a135	MKYKRIVFKVGTSSITHSDGSLSRGKIQTITRQLAALHHAGHELVLVSSGAVAAGFGALG					
	10	20	30	40	50	60
m135.pep	FKKRPVKIADKQASAAVGGQLLMEEYTANLSSDGIVSAQILLSRADFADKRRYQNAGGAL					
a135	FKKRPVKIADKQASAAVGGQLLMEEYTANLSSDGIVSAQILLSRADFADKRRYQNAGGAL					
	70	80	90	100	110	120
m135.pep	SVLLQRRVAVPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGNPNS					
a135	SVLLQRRVAVPIINENDTVSVEELKIGDNDTLSAQVAAMIQADLLVLLTDIDGLYTGNPNS					
	130	140	150	160	170	180
m135.pep	NPDAVRLDKIEHNHEIEMAGGSGSANGTGMLTKIKAATIAAESGVPVYICSSLKPDA					
a135	NPDAVRLDKIEHNHEIEMAGGSGSANGTGMLTKIKAATIAATESGVPVYICSSLKPDA					
	190	200	210	220	230	240

385

	250	260	270	280	290	300
m135.pep	LAEEAAEHQADGSFFVPRAKGLRTQKQWLAFYSESRGSVYVDEGAEHALSEQGKSLMSGI					
a135	LAEEAADNQADGSFFVPRAKGLRTQKQWLAFYSESRGGVYVDEGAEHALSEQGKSLMSGI					
	250	260	270	280	290	300
	310	320	330	340	350	360
m135.pep	AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAEDLLKSRKAKGVFIHRDDWISITP					
a135	AGIEGHFSRMDTVTVYSKATKQPLGKGRVLFSGAAEDLLKLRKAKGVFIHRDDWISITP					
	310	320	330	340	350	360
	370					
m135.pep	EIRLLLTEFX					
a135	EIRLLLTEFX					
	370					

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 523>:

```

g136.seq
1  ATGGAAATCC GGTTCAGAC AGCATTTTTC CGTTTGGTTC AGatgaAAAC
51  AAACGCTtca aTTCTtaccg caACACGCCT TGTATTTCCT GccgCTGCCG
101 CACGGACAGG GATCGTTCCT GCCGgtTTTT TCCCCTTCCC TGCGGACGGT
151 TTGCGGTTTG TTGATGACCG CCTGCCAGTA GCGGTAGATG TCtgccagcg
201 cgTAAGGCag tTCGGAcgca agttccgcca gctcgccctc ggTGAATTGC
251 AGgcgataa cgccggttTC CTCTTCGTCg taaatgccgc ccactgccat
301 cagcGGGTAA AACAGCTCTT CAAACGCTTC ATCATCGGCG GCTTCAAACC
351 AATCGGTCGG CACAATGTCC AAACCGTAAA GATAGGCGTT GCACCAAGTG
401 TAAAAATCGC TGCCGCCCTC GCCGTCGTCG TAGAGCCACA AATCGGCAG
451 CTTTTTATCC GACATCGCGG CGGTTGTTTC CATCGCCATT GCCAAAACCA
501 GCCGTTTCGAT TTCGGAACGT TCGGCGGCGG TAAATGCGA TTCGTCGCCC
551 AACACTTCGG GCAGCCAGTC GAGCGGTGCC AATTGTCCG GCCCGCTCAA
601 CAGCGCCGTC ATAAAACCTT GAACCTCGTC GCAACGCATC GTGTTGCCTT
651 GTTCGCTTTT GGCATCCAAT AA

```

This corresponds to the amino acid sequence <SEQ ID 524; ORF 136.ng>:

```

g136.pep
1  MEIRFQTAFL RLVQMKTNAS ILTATRLVFP AAAARTGIVP AGFFPFPADG
51  LRFVDDRLPV AVDVCQVRVQ FGRKFRQLAF GELQADNAVF LFFVNAAHCH
101 HGVKQLFKRF IIGGFKPIGR HNVQTVKIGV APSVKIAAAL AVVVEPQIGQ
151 LFIHRHGGCF HRHCQNQPF DFGTFGGGKLR FVAQHFGQPV ERCQFVRPAQ
201 QRRHKTLLNLV ATRHVALFAF GIQ*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 525>:

```

m136.seq
1  ATGGAAACAA ACGCTTCAAT TCTTACCGCA ACACGCCTTG TATTTTCTGC
51  CGCTGCCGCA CGGACAGGGA TCGTTCCTGC CTGTTTTTTC GCCTTCCCTG
101 CGGACGGTTT GCGGTTTGTT GATGACTGCC TGCCAGTAGC GGTAGATATC
151 CGCCAATGCA TAAGGCAACT CGGATTCCAG TTCCGCCAGC TCGCCTTCTG
201 TGAATTGCAG ACGGATAGCG CCGTTTTCCT CTTGTCGTA AATACCGCCC
251 AATGCCATGA TGGGATAAAA CAACTCTTCA AACGCTTCAT CATCGACGGC
301 TTCAAACCAA TCGGTCGGCA CAATATCCAA ACCGTAAAGA TAAGCATTGC
351 ACCATGTGTA AAAATCGCTG CCGCCGTCTT CGTTTTATA CAGCCACAAA
401 TCGGGCAGTT TTTTATCCGA CATCGCGGCG GTTGTTTCCA TCGCCATTGC
451 CAAAACCAGC CGTTCGATTT CGGAACGTTT GCGGCGGTA AATTGCGATT
501 CGTCGCCCAA CACTTCGGGC AGCCAGTCGA GCGGTGTCAA TTTGTCCGGC
551 CCGCTCAACA GCGCCGTCAT AAAACCTTGA ACCTCGTCGC AACGCATCGT
601 GTTGCCCTGT TCGCTTTTGG CATCCAACAA TTCGCTCAAC CGCCGTTTGG
651 ATGCTTCGGT AAATTTTCGG GAATCCATCA TTTTCTTTT CAAATGGGTT
701 TTGCGCCCTA TTATCGCCGC AATGCCGTCT GA

```

This corresponds to the amino acid sequence <SEQ ID 526; ORF 136>:

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 136 shows 85.6% identity over a 209 aa overlap with a predicted ORF (ORF 136.ng) from *N. gonorrhoeae*:

m136/g136

```

m136.pep      10          20          30          40
                METNASILTATRLVFSAAAARTGIVPACFFAFADGLRFVDCLPV
                |:|||||
g136           MEIRFQTAFLRLVQMKTNASILTATRLVFPAARAGTGVPPADGLRFVDDRLPV
                10          20          30          40          50          60

                50          60          70          80          90          100
m136.pep      AVDIRQCIRQLGFQRQLAFCELOQDSAVFLFVVNTAQCHDGIKQLFKRFIIDGFKPIGR
                |||: | :||: | ||||| ||:|:|||||:|:| | :||| |||||
g136           ADVCQVRQFGRKFRQLAFGELOQDNAVFLFVNAAHCHHGKQLFKRFIIIGGFKPIGR
                70          80          90          100          110          120

                110         120         130         140         150         160
m136.pep      HNIQTVKISIAPCVKIAAAVFVIQPQIGOFFIRHRGGCFHRHCQNQPFDFGTFGGKKLR
                ||:| ||||:|:| ||||| :|:|||||:||||| ||||| |||||
g136           HNVTQTKIGVAPSVKIAAALAVVVEPQIGQLFIRHRGGCFHRHCQNQPFDFGTFGGKKLR
                130         140         150         160         170         180

                170         180         190         200         210         220
m136.pep      FVAHQFGQPVERCQFVRPAQQRRHKTLNLVATHRVALFAFGIQQAQPPFCGKFSGIH
                ||||| ||||| ||||| ||||| |||||
g136           FVAHQFGQPVERCQFVRPAQQRRHKTLNLVATHRVALFAFGIQX
                190         200         210         220

                230         240
m136.pep      HFPPOMGFAPYYRRNAVX

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 527>:

a136.seq

1	ATGGAACAA	ACGCTTCAAT	TCTTACCGCA	ACACGCCTTG	TATTTTCTGC
51	CGCTGCCGCA	CGGACAGGGA	TCGTTCTCTG	CTGTTTTTTC	GCCTTCCCTG
101	CGGACGGTTT	CGCGCTGTT	GATGACCGCC	TGCCAGTAGC	GGTAGATATC
151	CGCCAATGCA	TAAGGCAACT	CGGATTCCAG	TTCCGCAAGC	TGCCTTCTG
201	TGAATTGCAG	ACGGATAGTG	CCGTTGTCTT	CTTCGTCGTA	AATACCGCCC
251	AATGCCATCA	TGGGATAAAA	CAACTCTTCA	AACGCTTCAT	CATCGACGGC
301	TTCAAACCAA	TCGGTCGGCA	CAATATCCAA	ACCGTAAGA	TAAGCAATTGC
351	ACCAATGTGA	AAAATCGCTG	CCGCGTCTT	CGTTTTATA	CAGCCACAAA
401	TCGGGCAAGT	TTTTATCCGA	CATCGCGCGC	GTTGTTTCCA	TGCCAATTGC
451	CAAAACGACG	CGTTCGATTT	CGGAACGTTC	GGCGGCGGTA	AATTGCGATT
501	CGTCGCCCCA	CACCTTCGGG	AGCCAGTCGA	GCGGTGTCAA	TTTGTCCGGC
551	CCGCTCAACA	GCGCGGTCAT	AAAACCTTGA	ACCTCGTCGC	AACGCATCGT
601	GTTGCCTTGT	TCGCTTTTGG	CATCCAACAA	TTCGCTCAAC	CGCGGTTTGG
651	ATGCTTCGGT	AAATTTTCGG	GAATCCATCA	TTTTCTTTT	CCAATGGGTT
701	TTGCGCCCTA	TTATAGTGG	TTAAATTTAA	ATCAGGACAA	GGCGACGAAG
751	CCGCAGACAG	TACAAATAGT	ACGCAAGGCG	GAGGCAACGC	CGTACTGGTT
801	TAAATTTAAT	CCACTATATC	GCCGCAATGC	CGTCTGA	

This corresponds to the amino acid sequence <SEQ ID 528; ORF 136.a>:

a136.pep

387

```

1  METNASILTA  TRLVFSAAAA RTGIVPACFF AFPADGLRLV DDRLPVAVDI
51  RQCIRQLGFQ  FRQLAFCELO  TDSAVVLFV  NTAQCHDGIK  QLFKRFIIDG
101 FKPIGRHNIQ  TVKISIAPCV  KIAAAVFVEI  QPQIGQFFIR  HRGGCFHRHC
151 QNQPFDFGTF  GGGKLRFVAQ  HFGQPVERCQ  FVRPAQQRH  KTLNLVATHR
201 VALFAFGIQQ  FAQPPFGCFG  KFSGIHHFPP  PMGFAPYYSG  LNLNQDKATK
251 PQTQVIVRQG  EATPYWFKFN  PLYRRNAV*

```

m136/a136 98.3% identity in 238 aa overlap

```

          10      20      30      40      50      60
m136.pep  METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRFVDDCLPVAVDIRQCIRQLGFQ
          |||
a136      METNASILTATRLVFSAAAAARTGIVPACFFAFPADGLRLVDDRLPVAVDIRQCIRQLGFQ
          10      20      30      40      50      60

          70      80      90     100     110     120
m136.pep  FRQLAFCELOTD SAVVLFVVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV
          |||
a136      FRQLAFCELOTD SAVVLFVVNTAQCHDGIKQLFKRFIIDGFKPIGRHNIQTVKISIAPCV
          70      80      90     100     110     120

          130     140     150     160     170     180
m136.pep  KIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTGGGKLRFVAQHFGQPVERCQ
          |||
a136      KIAAAVFVFIQPQIGQFFIRHRGGCFHRHCQNQPFDFGTGGGKLRFVAQHFGQPVERCQ
          130     140     150     160     170     180

          190     200     210     220     230     240
m136.pep  FVRPAQQRHKT LNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPPMGFAPYYRR
          |||
a136      FVRPAQQRHKT LNLVATHRVALFAFGIQQFAQPPFGCFGKFSGIHHFPPMGFAPYYSG
          190     200     210     220     230     240

m136.pep  NAVX
a136      LNLNQDKATKPQTQVIVRQGEATPYWFKFNPLYRRNAVX
          250     260     270

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 529>:

```

g137.seq
1  ATGATTATCC ATCACcaatT CGATCCCCTC CTATCAGTA TCGGCCCGCT
51  TGCCGTCGCG TGGTATGCCT TAAGCTACAT CCTCGGATTT ATTCTTTTTA
101 CCTTTCTCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCTG TTTTACCAAA
151 GAATCGCTCG ACGACTTCCT GACATGGGGC ATTTTGGGCG TGATTTTGGG
201 CGGACGCTTG GGCTATGTCC TGTTTTACAA ATTCTCCGAC TACCTCGCCC
251 ATCCGCTTGA TATTTTCAAG GTATGGGAAG GCGGAATGTC GTTCCACGGC
301 GGCTTTTGGG GTGTAGTTAT TGCCATATGG TTGTTTCAAG GCAAGCACGG
351 CATCGGCTTC CTCAAAGTGA TGGACACGGT CGCGCCGCTC GTTCCGCTGG
401 GTCTCGCTTC GGGACGTATC GGCAACTTTA TCAACGGCGA ACTTTGGGGA
451 CGCATTACCG ACATTAACGC ATTTTGGGCA ATGGGCTTCC CGCAAGCGCA
501 TTACGAAGAT GCCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAATGGC
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCCCTTGAAG GCATCTGCCT GTTCGCGCTC GTTTGGCTGT TTTCCAAAAA
651 ACCGCGCCCG ACCGGGCGA CTGCCGCGCT TTTTCTCGGC GGCTACGGCG
701 TGTTCCGCTT TATTGCCGAA TTGCGCGGCC AACCCGACGA CTATCTCGGG
751 CTGCTGACCT TGGGGCTGTC GATGGGGCAA TGGTTGAGCG TCCCGATGAT
801 TGTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AACAGCACT
851 GA

```

This corresponds to the amino acid sequence <SEQ ID 530; ORF 137.ng>:

g137.pep

m137.seq

m137.pcp

m137/g137

	10	20	30	40	50	60
m137.pep	MITHPQFD	PVLISIGPLAVRWY	ALSYILGFIL	FTFLGRRRIA	QGLSVFTKES	LDDFLTGW
g137	MIHHQFD	PVLISIGPLAVRWY	ALSYILGFIL	FTFLGRRRIA	QGLSVFTKES	LDDFLTGW
	10	20	30	40	50	60
	70	80	90	100	110	120
m137.pep	ILGVILGGR	LGVLFYKFS	DYLAHPLD	IFKWEGGMS	FHGGFLGV	VIARLFGRKHGIGF
g137	ILGVILGGR	LGVLFYKFS	DYLAHPLD	IFKWEGGMS	FHGGFLGV	VIARLFGRKHGIGF
	70	80	90	100	110	120
	130	140	150	160	170	180
m137.pep	LKLMDTVAP	LVPLGLASGR	IGNFINGEL	WGRVTDINA	FWAMGFPQ	ARYEDAEAAHNPLW
g137	LKLMDTVAP	LVPLGLASGR	IGNFINGEL	WGRITDINA	FWAMGFPQ	AHYEDAEAAHNPLW
	130	140	150	160	170	180
	190	200	210	220	230	240
m137.pep	AEWLQQYGM	LPRHPSQLYQ	FALEGICLF	TVIWLFSKK	KORSTGOVA	SLFLGGYGI

g137	AEWLQQYGMLPRHPSQLYQFALEGICLFAVVVWLSFKKPRPTGQTAAFLGGYGVFRFIAE	190	200	210	220	230	240
		250	260	270	280		
m137.pep	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX						
g137	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX	250	260	270	280		

```

a137.seq
1  ATGATTACCC ATCCCCAATT CGACCCCGTC CTTATCAGTA TCGGCCCGCT
51  TGCCGTCGCG TGGTATGCCC TAAGCTACAT CCTCGGATTT ATTCTTTTTT
101 CTTTCTCCGG CAGAAGGCGC ATCGCGCAAG GCTTGTCCGT TTTTACCAAA
151 GAATCGTCTG ACGACTTCCT GACATGGGGC ATTTTGGGCG TAATTTTGGG
201 CGCGCGTTTG GGTTAGCTGC GTTTTACAA GTTTTCCGAC TACCTCGCCC
251 ATCCGCTGTA TATTTTCAAG GTATGGGAAG CGGGAATGTC GTTCCACGGC
301 GGCTTTTGG GTGTAGTTAT TGCCATATGG TTGTTCGGTC GCAAACACGG
351 CATCGGCTTC CTCAAAGTCA TGGACACGGT CGCACCGCTC GTTCCACTGG
401 GTCTCGCTTC GGGACGTATC GGAACCTTCA TCAACGGCGA ACTTTGGGGA
451 CGCGTTACCG ACATCAACGC ATTTTGGGCA ATGGGCTTCC CGCAGGCGCG
501 TTACGAAGAC CTCGAAGCCG CCGCGCACAA TCCGCTTTGG GCAGAAATGG
551 TGCAACAATA CGGTATGCTG CCGCGTCATC CCTCGCAGCT TTATCAGTTT
601 GCATCTGAAG GCATCTGCCT GTTCCGCGTC GTTTGGCTGT TCTCTAAAAA
651 ACAGCGGCCG ACATCGCAAG TCGCCTCACT CTTCCTCGGC GGTCTAGGGA
701 TATTCGCTTT CATTGCCGAA TTTGCACGCC AATCCGACGA CTATCTCGGG
751 CTGCTGACCT TGGGCTGTC GATGGGGCAA TGGTGAGCG TCCCGATGAT
801 TGTTTTGGGT ATCGTCGGCT TTGTCCGGTT CGGCATGAAA AAACAGCACT
851 GA

```

a137.pep

1	MITHPQFDPV	LISIGPLAVR	WYALSYLGF	ILFTFLGRRR	IAQGLSVFTK
51	ESLDDFLTWG	ILGVILGGRL	GYVLFYKFSD	YLAHPLDIFK	VWEGGMSFHG
101	GFLGVVIAIW	LFGKRKHGIGF	LKLMDTVAPL	VPLGLASGRI	GNFINGELWG
151	RVTDINAFWA	MGPFQARYED	LEAAAHNPLW	AEWLQQYQML	PRHPSQLYQF
201	<u>ALLEGICLFAV</u>	<u>WVLFSSKKQRP</u>	<u>TGQVASLFLG</u>	<u>GYGIFRFIAE</u>	<u>FARQPPDYLG</u>
251	<u>LLTLLGLSMQG</u>	<u>WLSVPMIVLG</u>	<u>IVGFRFGMK</u>	<u>KQH*</u>	

	10	20	30	40	50	60
m137.pep	MITHPQFDPVLISIGPLAVRWYALS	YILGFILFTFLGRRRIAQGLSVFTKESLDDFLT	WG			
a137	MITHPQFDPVLISIGPLAVRWYALS	YILGFILFTFLGRRRIAQGLSVFTKESLDDFLT	WG			
	10	20	30	40	50	60
m137.pep	ILGVILGGRLG	YVLFYKFS	DYLAHPLDIFK	VWEGGMSFHGGF	LG	GVVIAIRLFGRKHGIGF
a137	ILGVILGGRLG	YVLFYKFS	DYLAHPLDIFK	VWEGGMSFHGGF	LG	GVVIAIRLFGRKHGIGF
	70	80	90	100	110	120
m137.pep	LKLM	DTVAPLVPLGLASGRIGNFINGELWGRVTD	DINAFWAMGFPQARYEDAEAAAHNPLW			
a137	LKLM	DTVAPLVPLGLASGRIGNFINGELWGRVTD	DINAFWAMGFPQARYEDAEAAAHNPLW			
	130	140	150	160	170	180
m137.pep	AEWLQ	QYGM	LP	HRHPSQLYQFALEGICLFTVIWLF	SKKQ	RSTGQVASLFLGGYGIFRFIAE
a137	AEWLQ	QYGM	LP	HRHPSQLYQFALEGICLFTVIWLF	SKKQ	RSTGQVASLFLGGYGIFRFIAE
	190	200	210	220	230	240

390

	190	200	210	220	230	240
	250	260	270	280		
m137.pep	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX					
a137	FARQPDDYLGLLTLGLSMGQWLSVPMIVLGIVGFVRFGMKKQHX					
	250	260	270	280		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 535>:

g138.seq

```

1  ATGGAGTTTG AAAACATTAT TTCCGCCGCC gaCAAGGCGC GTATCCTTGC
51  CGAAGCACTG CCTTACatcc gccgGTTTTT CCGTTCGGTC GCCGTCATCA
101 AGTATGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
251 GCGAATTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGAC GATGGATATT
301 GTCGAAATGG TATTGGGCGG GCACGTCAAC AAGGAAATCG TGTCGATGAT
351 TAACACATAT GGAGGGCACG CGGTCGGCGT GAGCGGGCGC GACGACCATT
401 TCATTAAGGC GAAGAAACTT TTGGTCGATA CGCCCGAACA GAATAGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGCGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT GGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAAATC TTGATGATGA CGAAtatcgc
651 cgGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACCT acgCCGAAAC
701 GGATTGATGG GCTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCcgcgcgc aACGGTGTGA AAGCCACGCA
801 CATCATCGAC GGCAGGTTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGGTCGATG ATTTTAGGCA GAGGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 536; ORF 138.ng>:

g138.pep

```

1  MEFENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
51  RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVDKETMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIAKAKL LVDTPQNSV
151 DIGQVGTVES IDTLVKGLI ERGCI PVVAP VGVGEKGEAF NINADLVAGK
201 LAEELNAEKL LMTNIAAGV DKTGNLLTKL TPKRIDGLIA DGTLYGGMLP
251 KIASAVEAAV NGVKATHIID GRPLNALLE IFTDAGIGSM ILGRGEDA*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 537>:

m138.seq

```

1  ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
51  CGAAGCGCTG CCTTACATCC GCCGGTTTTT CCGTTCGGTC GCCGTCATCA
101 AATACGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
251 GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
301 GTCGAAATGG TGTGGGCGG GCATGTCAAT AAAGAAATCG TGTCGATGAT
351 TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
401 TCATTAAGGC GAAGAAACTT TTGATCGATA CGCCCGAACA GAATGGCGTG
451 GACATCGGAC AGGTCGGTAC GGTGGAAGC ATCGATACCG GTTTGGTTAA
501 AGGGCTGATA GAACGTGGCT GCATTCCCGT CGTCGCCCCC GTCGGCGTAG
551 GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT AGCAGGCAAA
601 TTGGCGGAAG AATTGAACGC CGAAAAATC TTGATGATGA CGAATATCGC
651 CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACCT ACGCCGAAAC
701 GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
751 AAAATCGCTT CTGCGGTCGA AGCCGCCGTC AACGGTGTGA AAGCCACGCA
801 TATCATCGAC GGCAGGTTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
851 ATGCCGGTAT CGGTTCGATG ATTTGGGCG GTGGGAAGA TGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 538; ORF 138>:

m138.pep

```

1  MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA

```

51 RDVVLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVTDKAMDI
 101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
 151 DIGQVGTVES IDTGLVKGLI ERGCIPVVAP VGVGEKGEAF NINADLVAGK
 201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
 251 KIASAVEAAV NGVKATHIID GRLPNALLLE IFTDAGIGSM ILGGGEDA*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

ORF 138 shows 98.0% identity over a 298 aa overlap with a predicted ORF (ORF 138.ng)

from *N. gonorrhoeae*:

m138/g138

	10	20	30	40	50	60
m138.pep	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
g138	MEFENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m138.pep	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKAMDI VEMVLGGHVNKEIVSMINTY					
g138	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVTDKETMDI VEMVLGGHVNKEIVSMINTY					
	70	80	90	100	110	120
	130	140	150	160	170	180
m138.pep	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
g138	GGHAVGVSGRDDHFIKAKKLLVDTPEQNSVDIGQVGTVESIDTGLVKGLIERGCIPVVAP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m138.pep	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
g138	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDGLIA					
	190	200	210	220	230	240
	250	260	270	280	290	299
m138.pep	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGGGEDAX					
g138	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLLEIFTDAGIGSMILGRGEDAX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 539>:

a138.seq

1 ATGGAGTCTG AAAACATTAT TTCCGCCGCC GACAAGGCGC GTATCCTTGC
 51 CGAAGCGCTG CCTTACATCC GCCGGTTTTC CGGTTCCGTC GCCGTCATCA
 101 AATACGGCGG CAACGCGATG ACCGAACCTG CCTTGAAAGA AGGGTTTGCC
 151 CGCGATGTCG TGCTGCTGAA GCTGGTCGGC ATTCATCCCG TCATCGTTCA
 201 CGGCGGCGGG CCGCAGATCA ATGCGATGCT TGAAAAAGTC GGCAAAAAGG
 251 GTGAGTTTGT CCAAGGAATG CGCGTTACCG ACAAAGAGGC GATGGATATT
 301 GTCGAAATGG TGTTGGGCGG GCATGTCAAT AAAGAAATCG TGTGATGAT
 351 TAACACATAT GGCGGACACG CGGTCGGCGT AAGCGGACGC GACGACCATT
 401 TCATTAAGGC GAAGAAACTT TTGATCGATA CGCCGAACA GAATGGCGTG
 451 GACATCGGAC AGGTCGGTAC GGTGGAAAGC ATCGATACCG GTTGGTTAA
 501 AGGGCTGATA GAACGTGGCT GCATCCCCGT CGTCGCCCCC GTCGGCGTAG
 551 GTGAAAAAGG CGAAGCGTTC AACATCAACG CCGATTGGT AGCAGGCAAA
 601 TTGGCGGAAG AATTGAACGC CGAAAACTC TTGATGATGA CGAATATCGC
 651 CGGTGTGATG GACAAAACGG GCAATCTGCT GACCAAACTC ACGCCGAAC
 701 GGATTGATGA ACTGATTGCC GACGGCACGC TGTATGGCGG TATGCTGCCG
 751 AAAATCGCTT CTGCGGTCGA AGCCGCGGTC AACGGCGTGA AAGCCACGCA
 801 TATCATCGAC GGCAGGGTGC CCAACGCGCT TTTGCTGGAA ATCTTTACCG
 851 ATGCCGGTAT CGGTTTCGATG ATTTTGGGCG GTGGGAAGA TGCCTGA

This corresponds to the amino acid sequence <SEQ ID 540; ORF 138.a>:

```
a138.pep
1  MESENIISAA DKARILAEAL PYIRRFSGSV AVIKYGGNAM TEPALKEGFA
51  RDVVLLKLVG IHPVIVHGGG PQINAMLEKV GKKGEFVQGM RVDKEAMDI
101 VEMVLGGHVN KEIVSMINTY GGHAVGVSGR DDHFIKAKKL LIDTPEQNGV
151 DIGQVGTVES IDTGLVKGLI ERGCI PVVAP VGVGEKGEAF NINADLVAGK
201 LAEELNAEKL LMMTNIAGVM DKTGNLLTKL TPKRIDELIA DGTLYGGMLP
251 KIASAVEAAV NGVKATHIID GRVPNALLE IFTDAGIGSM ILGGGEDA*
```

m138/a138 99.7% identity in 298 aa overlap

	10	20	30	40	50	60
m138.pep	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
a138	MESENIISAADKARILAEALPYIRRFSGSVAVIKYGGNAMTEPALKEGFARDVVLLKLVG					
	10	20	30	40	50	60
	70	80	90	100	110	120
m138.pep	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVDKEAMDIVEMVLGGHVNKEIVSMINTY					
a138	IHPVIVHGGGPQINAMLEKVGKKGEFVQGM RVDKEAMDIVEMVLGGHVNKEIVSMINTY					
	70	80	90	100	110	120
	130	140	150	160	170	180
m138.pep	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCI PVVAP					
a138	GGHAVGVSGRDDHFIKAKKLLIDTPEQNGVDIGQVGTVESIDTGLVKGLIERGCI PVVAP					
	130	140	150	160	170	180
	190	200	210	220	230	240
m138.pep	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
a138	VGVGEKGEAFNINADLVAGKLAELNAEKLMMTNIAGVMDKTGNLLTKLTPKRIDELIA					
	190	200	210	220	230	240
	250	260	270	280	290	299
m138.pep	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRLPNALLEIFTDAGIGSMILGGGEDAX					
a138	DGTLYGGMLPKIASAVEAAVNGVKATHIIDGRVPNALLEIFTDAGIGSMILGGGEDAX					
	250	260	270	280	290	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 541>:

```
g139.seq
1  ATGCGAACCA CCTCAACCTT CCCTACAAAA ACTTTCAAAC CGGCTGCCAT
51  GGCCTTAGCT GTTGCAACAA CACTTTCTGC CTGCTTAagg ggcggcggag
101 gcGGCACTTC TGCTCCCGAC TTTAATGCAG GCGGCACCGG TATCGGCAGC
151 AACAGCAGGG CAACGATAGC GGAATCAGCA GCAGTATCTT ACGCCGGTAT
201 AAAAAACGAA ATGTGCAAAG ACAGAAGCAT GCTCTGTGCC GGTCCGGGATG
251 ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAAAGCCCC CCGAATCTGC
301 ATACCGGAGA CTTTTCAAAC CCAATGACC AATATTAAGA ATATGATCAA
351 CCTCAAAACCT GCAATTGAAG CAGGCTATAC AGGACGCGGG GTAGAGGTAG
401 GTATCGTCEA TACAGGCGAA TCCGTCGCA GCATATCCTT TCCCGAATCG
451 TATGGCAGAA AAGAACACGG CTATAACGAA AATTACAAAA ACAAATTACA
501 AAAACTATAC GGCGTATATG CGGAAGGAAG CGCCTGA
```

This corresponds to the amino acid sequence <SEQ ID 542; ORF 138.ng>:

```
g139.pep
1  MRTTSTFPTK TFKPAAMALA VATTLSACLG GGGGTSAPD FNAGGTGIGS
51  NSRATIAESA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKIKAPRIC
101 IPETFQTQMT NIKNMINKLP AIEAGYTGRG VEVGIVDTGE SVGSISFPPEL
151 YGRKEHGYNE NYKNKLQKLY GVYAEGSA*
```

```
m139.seq
1  ATGCGAACGA CCCCAACCTT CCCTACAAAA ACTTTCAAAC CGACTGCCAT
51  GCGGTAGCT GTTGCAACAA CACTTCTGCG CTGCTTAGGC GGCGGCGGAG
101 GCGGCACTTC TGCGCCCGCA TTCAATGCAG GCGGTACCGG TATCGGCAGC
151 AACAGCAGAG CAACAACAGC GAAATCAGCA CGACTATCTT ACGCGGGTAT
201 CAAGAACGAA ATGTGCAAAAG AAGAAGCAT GCTTGTGGT GCGTCGGGATG
251 ACGTTGCGGT TACAGACAGG GATGCCAAAA TCAATGCCCC CCCCCGAATC
301 TGCATACCGG AGACTTTCCA AACCCAAATG ACGCATtACA AGAATTTGAT
351 CAACCTCAAA CCTGCAATTG AAGCAGGCTA TACAGGA_GCG GGGGTAGAGG
401 TAGGTATCGT GCACACAGGC GAAATCCGTG GCAGCATATC CTTTCCCGAA
451 CTGTATGGCA GAAAAGAA CAAGCTATAC GAAAATTACG AAAAATATA
501 CGCGGTATAT GCGGAAGGAA GCGCTTGA
```

This corresponds to the amino acid sequence <SEQ ID 544; ORF 138>:

m139.pep

1	<u>MRTTPTFPTK</u>	<u>TFKPTAMALA</u>	VATTLSACLG	GGGGTSAPD	FNAGGTGIGS
51	NSRATTAKSA	AVSYAGIKNE	MCKDRSMLCA	GRDDVAVTDR	DAKINAPPRI
101	CIPETFQTQM	THYKNLNLK	PAIEAGYTGR	GVEVGIVDTG	ESVGSISFPE
151	LYGRKEHGYN	ENYEKLYGVY	AEGSA*		

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 138 shows 92.2% identity over a 179 aa overlap with a predicted ORF (ORF 138.ng) from *N. gonorrhoeae*:

```

m139/g139

      10          20          30          40          50          60
m139.pep  MRTTPTFPFTKTFKPTAMALAVATTLSACLGGGGGGTSAPDFNAGGTGIGSNSRATTAKSA
          ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : |||
g139      MRTTSTFPFTKTFKPAAMALAVATTLSACLGGGGGGTSAPDFNAGGTGIGSNSRATIAESA
          10          20          30          40          50          60

      70          80          90          100         110          120
m139.pep  AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKINAPPRICIPETFQTQMTHYKNLNLK
          ||| ||| ||| ||| ||| ||| ||| ||| : ||| ||| ||| ||| ||| : ||| |||
g139      AVSYAGIKNEMCKDRSMLCAGRDDVAVTDRDAKIKAP-RICIPETFQTQMTNIKNMINLK
          70          80          90          100         110

      130         140         150         160          170
m139.pep  PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENY----EKLYGVYAEGSAX
          ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| ||| ||| |||
g139      PAIEAGYTGRGVEVGIVDTGESVGSISFPELYGRKEHGYNENYKNLQKLYGVYAEGSAX
          120         130         140         150         160         170

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 545>:

```

a139.seq
1  ATGCGAACGA  CCCCAACCTT  CCCTACAAAA  ACTTTCAAAC  CGGCTGCCAT
51  GGGCTTAGCT  GTTGCACAA  CACTTTCTGC  CTGCTTAGGC  GGGCGCGGAG
101 GCGGCACTTC  TGCGCCGAC  TTCATGCGAG  CGGCGCACGG  TATCGGCAGC
151 AACAGCAGGG  CAACAACAGC  GAAATCAGCA  GCAATATCTT  ACGCCGGTAT
201 CAAGAACGAA  ATGTGCAAA  ACAGAAGCAT  GCTCTGTGCC  GGTGCGGATG
251 ACGTTGCGGT  TACAGACAGG  GATGCCAAAA  TCAATGCCCC  CCCCAGAACT
301 TGCATACCGG  AGACTTTACA  AACCCAAATG  ACGCAT. A  AGAATTTGAT
351 CAACCTCAA  CCTGCAATTG  AAGCAGGCTA  TACAGGACGC  GGGGTAGAGG
401 TAGGTATCGT  CGACACAGGC  GAATCCGTGC  GCAGCATATC  CTTTCCGAA
451 CTGTATGGCA  GAAAAGAACA  CGGCTATAAC  GAAAATTAC.  AAAAATATA
501 CGCGGTATAT  GCGGAAGGAA  GCGCCTGA

```

This corresponds to the amino acid sequence <SEQ ID 546; ORF 139.a>:

a139.pap
1 MRTTPTEPTK TEKPAAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS

394

51 NSRATTAKSA AISYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPRI
 101 CIPETLQTQM THXKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
 151 LYGRKEHGYN ENYXKLYGVY AEGSA*

m139/a139 97.1% identity in 175 aa overlap

	10	20	30	40	50	60
m139.pep	MRTTPTFTKTKFKPTAMALAVAT	TL	SACLG	GGGGG	TSAPDF	NAGGTGIGSNSRATTAKSA
a139	MRTTPTFTKTKFKPAAMALAVAT	TL	SACLG	GGGGG	TSAPDF	NAGGTGIGSNSRATTAKSA
	10	20	30	40	50	60
	70	80	90	100	110	120
m139.pep	AVSYAGIKNEMCKDRSMLCAGR	DDVAVTDR	DAKINAPP	RICIPET	FQTMTHYKNLINLK	
a139	AISYAGIKNEMCKDRSMLCAGR	DDVAVTDR	DAKINAPP	RICIPET	LQTQMTHXKNLINLK	
	70	80	90	100	110	120
	130	140	150	160	170	
m139.pep	PAIEAGYTGRGVEVGIVDTG	ESVGSISF	P	ELYGRKEHGYN	ENYKLYGVYAEGSAX	
a139	PAIEAGYTGRGVEVGIVDTG	ESVGSISF	P	ELYGRKEHGYN	ENYXKLYGVYAEGSAX	
	130	140	150	160	170	

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 547>:

g140.seq

```

1 Atgtcggcac gCGGCAAGGG GGCAGgctat ctcAACAGTA CCGGACGACa
51 TGTTCCTTC CTGAGTGCCG CCAAATCGG GCAGGATTAT TCTTCTTCA
101 AAAATATCAA AACCGACGGC GGTCTGCTGG CTTCCTCGA CAGCGTCGAA
151 AAAACAGCGG GCAGTGAAGG CGACACGCCG TCCTATTATG TCCGTCGCGG
201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCCGC CCCGCCGGTC
251 TGAAACACGC CGTAGAACAG GGCGGCAGCA ATCTGGA AAA CCTGATGGTC
301 GAGCTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAACTGC
351 GGTGCGCCGAC CGCACAGATA TGCCGGGCAT CCGCCTACGG CGCACAACTT
401 TCCGCACAGC GGCAGCCGTA CAGCATGCGA ATACCGCCGA CGGCGTACGc
451 aTCTTcaaCA GTCTCGCCGC TAccgTCTat GccgACAGTG CCGCCGCCCA
501 TGccgATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
551 ACAACGGTAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGA
601 ACGTGGGAAC AGGCGGGTGT CGAAGGCAAA ATGCGCGGCA GTACCCAAAC
651 TATCGGCATT GCCCGCAAAA CCGCGCAAAA TAGCAGACGA GCCGCCACAC
701 TGGGCATAGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC
751 GACAGCATTa GTCTGTTTGC AGGCATACGG CACGATGTGG GCGATATCGG
801 CTATCTCAAA GGCCTGTTCT CctaCGGACG CTACAAAAC AGCATCAGCC
851 GCAGCACCGG TGCAGATGAA TATGCGGAAG GCAGCGTCAA CGGCACGCTG
901 ATGCAGCTGG GCGCACTGGG TGGTGTCAAC GTTCCGTTG CCGCAACGGG
951 AGATTTGACG GTTGAAGGCG GTCTGCGCCA CGACCTGCTC AAACAGGATG
1001 CATTCGCCGA AAAAGGCagt GCTTTGGGCT GGAGCGGCAA CAGCCTCACT
1051 GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAAGTGTGCG AACCTTGAG
1101 CGATAAAGCC GTCCTGTCTG CGACGGCGGG CGTGGAACGC GACCTGAACG
1151 GACGCGACTA CGCGGTAACG GCGGGCTTTA CCGGCGCGGC TGCAGCAACC
1201 GGCAAGACCG GTGCACGCAA TATGCCGCAC ACCCGCCGGG TTGCCGGTCT
1251 GGGGGTGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301 GCTACACCGG TTCAAACAG TACGGCAACC ACAGCGGACA AATCGGCGTA
1351 GGCTACCGGT TCTGA
```

This corresponds to the amino acid sequence <SEQ ID 548; ORF 140.ng>:

g140.pep

```

1 MSARGKGAGY LNSTGRHVFP LSAAKIGQDY SFFKNIKTDG GLLASLDSVE
51 KTAGSEGDTP SYVVRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
101 ELDAESSAT PETVETAVAD RTDMPGIRLR RTTFRATAAV QHANTADGVR
151 IFNSLAATVY ADSAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQDDGG
201 TWEQGGVEGK MRGSTQTIGI AAKTGENTTA AATLGIGRST WSENSANAKT
```


395

251 DSISLFAIR HDVGDIGYLK GLFSYGRYKN SISRSTGADE YAEGSVNGTL
 301 MQLGALGGVN VPFAATGDLT VEGGLRHDLL KQDAFAEKGS ALGWSGNSLT
 351 EGTLVGLAGL KLSQPLSDKA VLSATAGVER DLNGRDYAVT GGFTGAAAT
 401 GKTGARNMPH TRRVAGLGVD VEFNGWNGL ARYSYTGSKQ YGNHSGQIGV
 451 GYRF*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 549>:

m140.seq
 1 ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG
 51 TGTTCCTTC CTGAGTGCCG CCAAATCGG GCAGGATTAT TCTTCTTCA
 101 CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCTCGA CAGCGTCGAA
 151 AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGCGG
 201 CAATGCGGCA CCGACTGCTT CCGCAGCGGC ACATTCCGCG CCCGCCGGTC
 251 TGAACACAGC CGTAGAACAG GCGGCGAGCA ATCTGGAAAA CCTGATGCTC
 301 GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
 351 GGCAGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCCTAC GGCGCAACTT
 401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CGGTGTACGC
 451 ATCTTCAACA GTCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA
 501 TGCCGATATG CAGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
 551 ACAACGGCAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGA
 601 ACGTGGGAAC AGGGCGGTGT TGAAGGCAAA ATGCGCGGCA GTACCCAAAC
 651 CGTCGGCATT GCCGCGAAAA CCGCGAAAA TACGACAGCA GCCGCCACAC
 701 TGGGCATGGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC
 751 GACAGCATTG GTCTGTTTGC AGGCATACGG CAGCATGCGG GCGATATCGG
 801 CTATCTCAAA GGCCTGTTCT CCTACGGACG CTACAAAAAC AGCATCAGCC
 851 GCAGCACCGG TCGGACGAA CATGCGGAAG GCAGCGTCAA CCGCACGCTG
 901 ATGCAGCTGG GCGCACTGGG CGGTGTCAAC GTCCGTTTG CCGCAACGGG
 951 AGATTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AAACAGGATG
 1001 CATTCGCCGA AAAAGGCAGT GCTTGGGCT GGAGCGGCAA CAGCCTCACT
 1051 GAAGGCACGC TGGTCGGACT CGCGGGTCTG AAGCTGTGCG AACCTTGAG
 1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
 1151 GACGCGACTA CACGGTAACG GGCGGCTTGA CCGGCGGAC TGCAGCAACC
 1201 GGCAAGACGG GGGCACGCAA TATGCCGCAC ACCCGTCTGG TTGCCGGCCT
 1251 GGGCGCGGAT GTCGAATTCT GCAACGGCTG GAACGGCTTG GCACGTTACA
 1301 GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
 1351 GGCTACCGGT TCTGA

This corresponds to the amino acid sequence <SEQ ID 550; ORF 140>:

m140.pep
 1 MSARGKGAGY LNSTGRRVPF LSAKIGQDY SFTNIETDG GLLASLDSVE
 51 KTAGSEDTL SYVRRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
 101 ELDAESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR
 151 IFNSLAATVY ADSTAAHADM QGRRLKAVSD GLDHNGTGLR VIAQTQDDGG
 201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT
 251 DSISLFAIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
 301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT
 351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
 401 GKTGARNMPH TRLVAGLGAD VEFNGWNGL ARYSYAGSKQ YGNHSGRVGV
 451 GYRF*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 140 shows 94.5% identity over a 454 aa overlap with a predicted ORF (ORF 140.ng) from *N. gonorrhoeae*:

m140/g140

	10	20	30	40	50	60
m140.pep	MSARGKGAGYLNSTGRRVPFLSAKIGQDYSFTNIETDGGLLASLDSVEKTAGSEDTL					
g140	MSARGKGAGYLNSTGRHVPFLSAKIGQDYSFFKNIKTDGGLLASLDSVEKTAGSEDTLP					
	10	20	30	40	50	60
	70	80	90	100	110	120

m140.pep	SYVVRGNAARTASAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAAAD
g140	SYVVRGNAARTASAAHSAPAGLKHAVEQGGSNLENLMVELDASESSATPETVETAVAD
	70 80 90 100 110 120
m140.pep	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNLSLAATVYADSTAAHADMQGRRLKAVSD
g140	RTDMPGIRLRRTTFRATAAVQHANTADGVRIFNLSLAATVYADSAAAHADMQGRRLKAVSD
	130 140 150 160 170 180
m140.pep	GLDHNGTGLRVIAQTQDDGGTWEQGGVEGKMRGSTQTIGIAAKTGENTTAAATLGMGRST
g140	GLDHNGTGLRVIAQTQDDGGTWEQGGVEGKMRGSTQTIGIAAKTGENTTAAATLGIGRST
	190 200 210 220 230 240
m140.pep	WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL
g140	WSENSANAKTDSISLFAGIRHDVGDIGYLKGLFSYGRYKNSISRSTGADEYAEGSVNGTL
	250 260 270 280 290 300
m140.pep	MQLGALGGVNVFFAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSLTGTLVGLAGL
g140	MQLGALGGVNVFFAATGDLTVEGGLRHDLLKQDAFAEKGSALGWSGNSLTGTLVGLAGL
	310 320 330 340 350 360
m140.pep	KLSQPLSDKAVLFATAGVERDLNGRDYTEVTGGFTGATAATGKTGARNMPHTRLVAGLGAD
g140	KLSQPLSDKAVLSATAGVERDLNGRDYAVTGGFTGAAAATGKTGARNMPHTRRVAGLGVD
	370 380 390 400 410 420
m140.pep	VEFGNGWNGLARYSYAGSKQYGNHSGRVGVGYRFX
g140	VEFGNGWNGLARYSYTGSKQYGNHSGQIGVGYRFX
	430 440 450

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 551>:

```

a140.seq
1  ATGTCGGCAG GCGGTAAGGG GGCAGGCTAT CTCAACCGTA CCGGACAACG
51  TGTTCCTTC CTGAGTGCCG CCAAAATCGG GCGGGATTAT TCTTTCTTCA
101 CAAACATCGA AACCGACGGC GGTCTGCTGG CTTCCCTCGA CAGCGTCGAA
151 AAAACAGCGG GTAGTGAAGG CGACACGCTG TCCTATTATG TCCGTGCGGG
201 CAATGCGGCA CGGACTGCTT CGGCAGCGGC ACATTCCGCG CCCGCCGGTC
251 TGAAACACGC CGTAGAACAG GCGGCGAGCA ATCTGGAAAA CCTGATGGTC
301 GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
351 GGCCGCCGAC CGCACAGATA TGCCGGGCAT CCGCCCCTAC GGCGCCAATT
401 TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCCGCCGA CGGTGTACGC
451 ATCTTCAACA ATCTCGCCGC TACCGTCTAT GCCGACAGTA CCGCCGCCCA
501 TGCCGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
551 ACAACGCTAC GGGTCTGCGC GTCATCGCGC AAACCCAACA GGACGGTGGA
601 ACGTGGGAAC AGGCGCGTGT TGAAGGCAAA ATGCGCGGCA GTACCCAAAC
651 CTGCGGCATT GCGCGGAAAA CCGCGGAAAA TACGACAGCA GCCGCCACAC
701 TGGGCATGGG ACACAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAACC
751 GACAGCATTA GTCTGTTTGC AGGCATACGG CACGATGCGG GCGATATCGG
801 CTATCTCAA GGCCTGTTCT CCTACGACG CTACAAAAAC AGCATCAGCC
851 GCAGCACCGG TCGGACGAA CATGCGGAAG GCAGCGTCAA CCGCACGCTG
901 ATGCAGCTGG GCGCACTGGG CGGTGTCAAC GTTCCGTTTG CCGCAACGGG

```

```

951 AGATTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AACAGGATG
1001 CATTGCGCGA AAAAGGCAGT GCTTTGGGCT GGAGCGGCAA CAGCATCACT
1051 GAAGGCACAC TGGTCGGACT CGCGGGTCTG AAGCTGTCGC AACCTTGAG
1101 CGATAAAGCC GTCCTGTTTG CAACGGCGGG CGTGGAACGC GACCTGAACG
1151 GACGCGACTA CACGGTAACG GCGGCTTTA CCGGCGCGAC TGCAGCAACC
1201 GGCAGACGG GGGCACGCAA TATGCCGCAC ACCCGCCTGG TTGCCGGTCT
1251 GGGCGCGGAT GTCGAATTCG GCAACGGCTG GAACGGCTTG GCACGTTACA
1301 GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
1351 GGCTACCGGT TCTGA

```

-- This corresponds to the amino acid sequence <SEQ ID 552; ORF 140.a>:

```

a140.pep
1 MSAGGKGAGY LNRTGQRPVF LSAKIGRDY SFTNIETDG GLLASLDSVE
51 KTAGSEGDTL SYYVRRGNAA RTASAAHSA PAGLKHAVEQ GGSNLENLMV
101 ELDAESSAT PETVETAAAD RTDMPGIRPY GATFRAAAV QHANAADGVR
151 IFNNLAATVY ADSTAAHADM QGRRLKAVSD GLDHNATGLR VIAQTQDGG
201 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGHST WSENSANAKT
251 DSISLFAGIR HDAGDIGYLK GLFSYGRYKN SISRSTGADE HAEGSVNGTL
301 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSIT
351 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT
401 GKTGARNMPH TRLVAGLGAD VEFNGWNL ARYSYAGSKQ YGNHSGRVGV
451 GYRF*

```

m140/a140 98.2% identity in 454 aa overlap

m140.pep	10	20	30	40	50	60
	MSARGKGAGYLNSTGRRVPFLSAKIGQDYSFTNIETDGGLLASLDSVEKTAGSEGDTL					
a140	MSAGGKGAGYLNRTGQRPVFLSAKIGRDYSFTNIETDGGLLASLDSVEKTAGSEGDTL					
	10	20	30	40	50	60
m140.pep	70	80	90	100	110	120
	SYYVRRGNAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDAESSATPETVETAAAD					
a140	SYYVRRGNAARTASAAHSAAPAGLKHAVEQGGSNLENLMVELDAESSATPETVETAAAD					
	70	80	90	100	110	120
m140.pep	130	140	150	160	170	180
	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNSLAATVYADSTAAHADMQGRRLKAVSD					
a140	RTDMPGIRPYGATFRAAAVQHANAADGVRIFNNLAATVYADSTAAHADMQGRRLKAVSD					
	130	140	150	160	170	180
m140.pep	190	200	210	220	230	240
	GLDHNGTGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGRST					
a140	GLDHNATGLRVIAQTQDGGTWEQGGVEGKMRGSTQTVGIAAKTGENTTAAATLGMGHST					
	190	200	210	220	230	240
m140.pep	250	260	270	280	290	300
	WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL					
a140	WSENSANAKTDSISLFAGIRHDAGDIGYLKGLFSYGRYKNSISRSTGADEHAEGSVNGTL					
	250	260	270	280	290	300
m140.pep	310	320	330	340	350	360
	MQLGALGGVNVFPAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSITTEGLTVGLAGL					
a140	MQLGALGGVNVFPAATGDLTVEGGLRYDLLKQDAFAEKGSALGWSGNSITEGLTVGLAGL					
	310	320	330	340	350	360
m140.pep	370	380	390	400	410	420
	KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD					
a140	KLSQPLSDKAVLFATAGVERDLNGRDYTVTGGFTGATAATGKTGARNMPHTRLVAGLGAD					

398

	370	380	390	400	410	420
	430	440	450			
m140.pep	VEFGNGWNGGLARYSYAGSKQYGNHSGRNVGVGYREFX					
a140	VEFGNGWNGGLARYSYAGSKQYGNHSGRNVGVGYREFX					
	430	440	450			

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 553>:

g141.seq

```

1  atgagcttca aAaccgATGC CGAAACCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTTGAC AACATTGAGC
101 CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAGCTG
151 CCGCAAAAAC AAGGCAGGCT GATTTTGGTT ACGCCATCA ACCCGACTCC
201 GCGGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
251 GCCATATCGG CAAAGACTCT GTGATTGCTT TGCGCGAGCC TTCTTTGGGT
301 CCGGTGTTTC GCGTGAAAGG CGGCGCGGCA GCGGCGGCT ACGCGCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGCGGAC TTCCACGCCA
401 TCGGTGCGGC GAATAACCTC CTCGCCGCCA TGCTCGACAA CCATATCTAC
451 CAAGGTAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GCGGCGCGCT
501 GGTGCATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGTATGGGCA
551 AGCCTGTtga cggCGTGATG CGtcccGACG GCTTCGACAT CACCGTCGCC
601 TCCGAAGTGa tggcggGTATT CTGCCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTtTt gGCAATATC TCGTCGCCTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTTGAAAG GCACACGGCG CGATGGCGGC ATTGCTAAAA
751 GATGCGATTA AGCCCAATTT GGTGCAAACC ATCGAAGGCA CTCCGGCCTT
801 TGACACGGC GGCCCGTTCG CCAACATCGC CCACGGCTGC AACTCCGTTA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG TCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CCTTGCCGGT TTGAAACCTG ATGCGGCAGT CGTCGTGGCG ACTGTCCGCG
1001 CCCTGAAATA CAACGGCGGC GTGGAACGCG CCAACCTTGG TGAAGAAAC
1051 CTCGAAGCCT TGGCAAAAGG TTGCCCAC CTGTTGAAAC ACATTTCAA
1101 CCTGAAAAAC GTATTCGGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1151 TGTCCGACTC CGATGCCGAG TTGGCGATGA TTGAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG CGGCGCGGG
1251 CCGCGCGGAT TTGGCGCGCA AAGTCGTCAA TGCCATCGAC AACCAACCTA
1301 ATAACCTCGG TTTCGCCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCAAAAAGT GTACGGCGCG GAAGATGTCG ATTTCAGCGC
1401 GGAAGCGTCT GCCGAAATCG CCTCGCTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATATT CATTGAGCGA CAACGCCAAA
1501 CTCTTGGGCT GCCCGAAGG CTTCCGCATC GCCGTACGCG GTATCACTGT
1551 TTCCGCGGCG CCGGGCTTCA TCGTTGCGTT GTGCGGCAAT ATGATGAAAA
1601 TGCCGGGCGT GCCGAAAGT CCGGTGCGG AGAAAATCGA TGTGGACGAA
1651 CACGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 554; ORF 141.ng>:

g141.pep

```

1  MSFKTDAETA QSSTMRPIGE IAAKLGLNVD NIEPYGHYKA KINPAEAFKL
51  PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFGVKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANL LAAMLNDHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPDVDM RPDGFEDITVA
201 SEVMAVFCLA KDISDLKERF GNILVAYAKD GSPVYAKDLK AHGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LEALAKGLPN LLKHISNLKN VFGLPVVVAL NRFVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAID NQPNNGFAY DVELGIKDKI
451 RAIQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPGEGFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDE
551 HGVIHGLF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 555>:

m141.seq

```

1  ATGAGCTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTTGGTCT GAATGCCGAC AACATTGAGC
101 CTTACGGTCA TTACAAGGCG AAAATCAATC CTGCCGAAGC GTTCAAACCTG
151 CCGCAAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC
201 GGCGGGCGAA GGCAAAACCA CCGTAACCAT CGGTTTGGCG GACGCGTTGC
251 GCCACATCGG CAAAGATGCC GTGATTGCCC TGCGCGAACC TTCTCTGGGG
301 CCGGTGTTCG GCGTGAAAGG CGGCGCGGCA GGCGGCGGCT ATGCCCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
401 TCGGTGCGGC AAATAATCTG CTGCGCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GGCGGCGCGT
501 GGTGCGATAT AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGTA
551 AACCCGTTGA CGGCGTGATG CGTCCTGACG GTTTCGATAT TACCGTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTTGAA
651 AGAGCGTTTG GGCAACATCC TTGTGCGCTA CGCCAAGAGC GGCAGCCCCG
701 TTTACGCCAA AGATTTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
751 GATGCGATTA AGCCCAACTT GGTGCAAACC ATCGAAGGCA CGCCCGCCTT
801 CGTACACGGC GGCCCGTTTC CCAACATCGC CCACGGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGGCG CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CCTTGCCGGT TTGAAACCTG ATGCGGCTGT TGTGCGGCG ACTGTCGCGC
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAT
1051 TTAGACGCTT TGGAAAAAGG TTTGCCCAAC CTGCTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTGCGAC TGCCCGTCGT CGTTGCGCTC AACCGCTTCG
1151 TGTCCGACGC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251 CGGCGCGGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
1301 ATAACCTCGG TTTGCGCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCCAAAAAGT GTACGGCGCG GAAGATGTTG ATTTCAGCGC
1401 GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501 CTGTTGGGCT GCCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCGT
1551 TTCCCGAGGC GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCCT GCCCAAAGTT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
1651 GAAGGCGTGA TTCACGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 556; ORF 141>:

m141.pep

```

1  MSFKTDAEIA QSSTMRPIGE IAAKLGLNAD NIEPYGHYKA KINPAEAFKL
51  PQKQGRLLIV TAINPTPAGE GKTTVTIGLA DALRHIGKDA VIALREPSLG
101 PVFGVKGGA GGYAQVLPD EDINLHFTGD FHAIGAANNL LAAMLNDNIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPDGVDM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKCRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRVSDADAE LAMIEKACAE
401 HGVEVSLTEV WGKGGAGGAD LARKVVNAIE SQTNNFGFAY DVELGIKDKI
451 RAIAQKVYGA EDVDFSABAS AEIASLEKLG LDKMPICMAK TQYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
551 BGVIHGLF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 141 shows 97.5% identity over a 558 aa overlap with a predicted ORF (ORF 141.ng) from *N. gonorrhoeae*:

m141/g141

	10	20	30	40	50	60
m141.pep	MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKL					
	PQKQGRLLIV					
g141	MSFKTDAEIAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKL					
	PQKQGRLLIV					
	10	20	30	40	50	60

400

m141.pep	70	80	90	100	110	120
	TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFGVKGGAAGGGYAQVLP					
g141	70	80	90	100	110	120
	TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFGVKGGAAGGGYAQVLP					
m141.pep	130	140	150	160	170	180
	EDINLHFTGDFHAIGAANNLLAAMLNDHIYQGNELNIDPKRVLWRRVVDMDRQLRNIID					
g141	130	140	150	160	170	180
	EDINLHFTGDFHAIGAANNLLAAMLNDHIYQGNELNIDPKRVLWRRVVDMDRQLRNIID					
m141.pep	190	200	210	220	230	240
	GMGKPVGDGVMRPGFDITVASEVMAVFCLAKDISDLKERLGNILVAYAKDGSFPVYAKDLK					
g141	190	200	210	220	230	240
	GMGKPVGDGVMRPGFDITVASEVMAVFCLAKDISDLKERFGNILVAYAKDGSFPVYAKDLK					
m141.pep	250	260	270	280	290	300
	ANGAMAALLKDAIKPNLVQTIETGPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
g141	250	260	270	280	290	300
	AHGAMAALLKDAIKPNLVQTIETGPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
m141.pep	310	320	330	340	350	360
	GFGADLGAEKFCDIKRLAGLKPDAVVVATVRALKYNGGVERANLGEENLDALEKGLPN					
g141	310	320	330	340	350	360
	GFGADLGAEKFCDIKRLAGLKPDAVVVATVRALKYNGGVERANLGEENLEALAKGLPN					
m141.pep	370	380	390	400	410	420
	LLKHISNLKNVFGLPVVVALNRFVSDADAEALAMIEKACAEHGVEVSLTEVWGKGAGGAD					
g141	370	380	390	400	410	420
	LLKHISNLKNVFGLPVVVALNRFVSDSDAEALAMIEKACAEHGVEVSLTEVWGKGAGGAD					
m141.pep	430	440	450	460	470	480
	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAIQKVYGAEDVDFSAEASAEIASLEKLG					
g141	430	440	450	460	470	480
	LARKVVNAIDNQPNFGFAYDVELGIKDKIRAIQKVYGAEDVDFSAEASAEIASLEKLG					
m141.pep	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
g141	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPGEGFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
m141.pep	550	559				
	PAAEKIDVDAEGVIHGLFX					
g141	PAAEKIDVDEHGVHGLFX					
	550					

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 557>:

```

a141.seq
1  ATGAGTTCA AAACCGATGC CGAAATCGCC CAATCCTCCA CCATGCGCCC
51  GATTGGCGAA ATTGCCGCCA AGCTGGGTTT GAACGTTGAC AACATTGAGC
101 CTTACGGTCA TTACAAAGCC AAAATCAATC CTGCCGAAGC GTTCAAAGT
151 CCGCAAAAAC AGGGCAGGCT GATTTTGGTT ACCGCCATCA ACCCGACTCC

```

```

201 GCGGGGCGAA GGTAACCA CCGTAACCAT CGGTTTGGCG GACGCATTGC
251 GCCATATCGG CAAAGACTCT GTGATTGCTT TCGCGAGACC TTCTTTGGGT
301 CCGGTGTTCG GCGTGAAAGG CGCGCGGCGA GCGGGCGGCT ATGCCCAAGT
351 TTTGCCGATG GAAGACATCA ACCTGCACTT CACCGGAGAT TTTCACGCCA
401 TCGGTGCGGC AAATAATCTG CTGCGCGCGA TGCTCGACAA CCATATCTAC
451 CAAGGCAACG AGTTGAACAT CGACCCCAAA CGCGTGCTGT GCGGCGCGCT
501 GGTCGATATG AACGACCGCC AGTTGCGCAA CATCATCGAC GGCATGGGCA
551 AGCCTGTTGA CGGCGTGATG CGTCTTGACG GTTTCGATAT TACCGTTGCT
601 TCCGAAGTGA TGGCGGTATT CTGTCTTGCC AAAGACATCA GCGATTGAA
651 AGAGCGTTTG GGCAACATCC TTGTCGCTTA CGCCAAAGAC GGCAGCCCCG
701 TTTACGCCAA AGATTGAAA GCGAATGGCG CGATGGCGGC ATTGCTTAAA
751 GATGCGATTA AGCCCAACTT GGTGCAAAAC ATCGAAGGCA CGCCCCGCTT
801 CGTACACGGC GGCCCGTTTC CCAACATCGC CCACGGCTGC AACTCCGTAA
851 CCGCAACCCG TCTGGCGAAA CACCTTGCCG ATTACGCCGT AACCGAAGCA
901 GGCTTCGCGC CGGACTTGGG CGCGGAAAAA TTCTGCGACA TCAAATGCCG
951 CTTGCGCGGT TTGAAACCTG ATGCGGCTGT TGTCTGGCGG ACTGTCGCGG
1001 CGTTGAAATA TAACGGCGGC GTGGAACGCG CCAACCTCGG CGAAGAAAAT
1051 TTAGACGCTT TGGAAAAAGG TTGCCCCAAC CTGCTGAAAC ACATTTCCAA
1101 CCTGAAAAAC GTATTGCGAC TGCCCGTCTG CGTTGCGCTC AACCGCTTCG
1151 TGTCCGACTC CGATGCCGAG TTGGCGATGA TTGAAAAAGC CTGTGCCGAA
1201 CACGGCGTTG AAGTTTCCCT GACCGAAGTG TGGGGCAAAG GTGGTGCGGG
1251 CGCGCGCGAT TTGGCGCGCA AAGTCGTCAA CGCCATTGAA AGTCAAACCA
1301 ATAACCTCGG TTTGCGCTAC GATGTCGAGT TGGGCATCAA AGACAAAATC
1351 CGTGCGATTG CCCAAAAGT GTACGGCGCG GAAGATGTTG ATTTACGCGC
1401 GGAAGCGTCT GCCGAAATCG CTTCACTGGA AAACTGGGC TTGGACAAAA
1451 TGCCGATCTG CATGGCGAAA ACCCAATACT CTTTGAGCGA CAACGCCAAA
1501 CTGTTGGGCT GCCCGAAGA CTTCCGCATC GCCGTGCGCG GCATCACCGT
1551 TTCCGAGGCG GCAGGTTTCA TCGTCGCCCT GTGCGGCAAC ATGATGAAAA
1601 TGCCCGGCCT GCCCAAAGT CCGGCTGCCG AGAAAATCGA TGTGGACGCA
1651 GAAGGCGTGA TTCACGGCTT GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 558; ORF 141.a>:

```

a141.pep
1 MSFKTDAEIA QSSTMRPIGE IAAKGLNVD NIEPYGHYKA KINPAEAFKL
51 PQKQGRILIV TAINPTPAGE GKTTVTIGLA DALRHIGKDS VIALREPSLG
101 PVFVGKGGAA GGGYAQVLPM EDINLHFTGD FHAIGAANNL LAAMLNDNHIY
151 QGNELNIDPK RVLWRRVVDN NDRQLRNIID GMGKPVGDVM RPDGFDITVA
201 SEVMAVFCLA KDISDLKERL GNILVAYAKD GSPVYAKDLK ANGAMAALLK
251 DAIKPNLVQT IEGTPAFVHG GPFANIAHGC NSVTATRLAK HLADYAVTEA
301 GFGADLGAEK FCDIKRLAG LKPDAAVVVA TVRALKYNGG VERANLGEEN
351 LDALEKGLPN LLKHISNLKN VFGLPVVVAL NRVSDSDAE LAMIEKACAE
401 HGVEVSLTEV WKGKGAGGAD LARKVNAIE SQTNNEGFAY DVELGKDKI
451 RAIQKVYGA EDVDFSAEAS AEIASLEKLG LDKMPICMAK TOYSLSDNAK
501 LLGCPEDFRI AVRGITVSAG AGFIVALCGN MMKMPGLPKV PAAEKIDVDA
551 EGVIHGLF*

```

m141/a141 99.5% identity in 558 aa overlap

```

          10      20      30      40      50      60
m141.pep  MSFKTDAEIAQSSTMRPIGEIAAKLGLNADNIEPYGHYKAKINPAEAFKL PQKQGRILIV
          |||
a141      MSFKTDAEIAQSSTMRPIGEIAAKLGLNVDNIEPYGHYKAKINPAEAFKL PQKQGRILIV
          10      20      30      40      50      60

          70      80      90      100     110     120
m141.pep  TAINPTPAGEGKTTVTIGLADALRHIGKDAVIALREPSLGPVFVGKGAAGGGYAQVLPM
          |||
a141      TAINPTPAGEGKTTVTIGLADALRHIGKDSVIALREPSLGPVFVGKGAAGGGYAQVLPM
          70      80      90      100     110     120

          130     140     150     160     170     180
m141.pep  EDINLHFTGDFHAI GAANNLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDNMNDRQLRNIID
          |||
a141      EDINLHFTGDFHAI GAANNLLAAMLNDNHIYQGNELNIDPKRVLWRRVVDNMNDRQLRNIID
          130     140     150     160     170     180

```

402

m141.pep	190	200	210	220	230	240
	GMGKPDVGVMRDPGFDITVASEVMAVFLAKDISDLKERLGNILVAYAKDGSFVYAKDLK					
a141	GMGKPDVGVMRDPGFDITVASEVMAVFLAKDISDLKERLGNILVAYAKDGSFVYAKDLK					
	190	200	210	220	230	240
m141.pep	250	260	270	280	290	300
	ANGAMAALLKDAIKPNLVQTIIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
a141	ANGAMAALLKDAIKPNLVQTIIEGTPAFVHGGPFANIAHGCNSVTATRLAKHLADYAVTEA					
	250	260	270	280	290	300
m141.pep	310	320	330	340	350	360
	GFGADLGAEKFCDIKRLAGLKPDAVVVATVRALKYNGGVERANLGEENLDALEKGLPN					
a141	GFGADLGAEKFCDIKRLAGLKPDAVVVATVRALKYNGGVERANLGEENLDALEKGLPN					
	310	320	330	340	350	360
m141.pep	370	380	390	400	410	420
	LLKHISNLKNVFLPVVVALNRFVSDADAELAMIEKACAEHGEVSLTEVWGKGGAGGAD					
a141	LLKHISNLKNVFLPVVVALNRFVSDADAELAMIEKACAEHGEVSLTEVWGKGGAGGAD					
	370	380	390	400	410	420
m141.pep	430	440	450	460	470	480
	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAIQAQVYGAEDVDFSAEASAEIASLEKLG					
a141	LARKVVNAIESQTNNFGFAYDVELGIKDKIRAIQAQVYGAEDVDFSAEASAEIASLEKLG					
	430	440	450	460	470	480
m141.pep	490	500	510	520	530	540
	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
a141	LDKMPICMAKTQYSLSDNAKLLGCPEDFRIAVRGITVSAGAGFIVALCGNMMKMPGLPKV					
	490	500	510	520	530	540
m141.pep	550	559				
	PAAEKIDVDAEGVIHGLFX					
a141	PAAEKIDVDAEGVIHGLFX					
	550					

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 559>:

```

g142.seq
1  ATGCGTGCCG  ATTTTCATGTT  TGCCGACAAT  ATGCCCGTGC  AGGTGCGCCA
51  ACGCGCCTTC  TATTTCAAGT  TGTCCCGTTT  TGCCGCGATG  CCAAATATGG
101 TAGGCAAAACC  GCTCTTCGGG  CGACAGGCCG  GTCAGCCCGG  CAAAATGTTC
151 GGCAACATCC  TGATGTTTCT  CCGCCAGCAT  ATTGATGCAG  AGGCTGCCGT
201 TTTCGACAG  GATcggaATG  AttcgCGCAC  TCCGTTTAT  GCACAGCATC
251 ACGGTCGGCG  GCTCGTCGGT  AACCGGCGCA  ACCGCCGTCA  TTGTAATGCC
301 GTAACGCCCT  GCCGCACCGT  CTGTCGTGAT  GACATGAACG  CCTGCCGCAC
351 AGGATGCCAT  CGCATCACGG  AACGAAGTTT  GAAAAGTTT  CTGCAAATCC
401 GCCATTTTTC  CCCTTTAAAC  CGTCCCCTAT  ATAAGAATGC  TGCACACAAG
451 GCATCCCCC  ATGTGCAGCA  GTTCTGA

```

This corresponds to the amino acid sequence <SEQ ID 560; ORF 142.ng>:

```

g142.pep
1  MRADFMFADN  MPVQVRQRAF  YFKLSRFAAM  PNMVGKPLFG  RQAGQPGKMF
51  GNILMFVRQH  IDAEAAVFRQ  DRNDSRTPVY  AQHHGRRLVG  NRRNRRHCNA
101 VTPCRTVCRD  DMNACRTGCH  RITERSLKSF  LQIRHFSPLN  RPLYKNAAHK
151 ASPHVQQF*

```


The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 561>:

```
m142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCCGTGC AGGTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC
151 GGCAACATCC TGATGTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCCGT AACCGGCGCG ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
351 AAGATGCCAT CGCATCACGG AACGAAGTTT GAAAATTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC TGTCCCTAT ATAAGAATGC TGCACACAAG
451 GCATCCCCcC ATGTGCAGCA GTTTTGA
```

This corresponds to the amino acid sequence <SEQ ID 562; ORF 142>:

```
m142.pep
1  MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF
51  GNILMFVRQR IDAAAVFRQ DRNDSRTPVD AOHGRRRLVG NRRDRRHCA
101 VTPCRTVCRD DMNACRARCH RITERSLKIF LQIRHFSPLN CPLYKNAAHK
151 ASPHVQOF*
```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 142 shows 93.7% identity over a 158 aa overlap with a predicted ORF (ORF 142.ng) from *N. gonorrhoeae*:

```
m142/g142
      10      20      30      40      50      60
m142.pep MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMFNGNLMFVRQR
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g142      MRADFMFADNMPVQVRQAFYFKLSRFAAMPNMVGKPLFGRQAGQPGKMFNGNLMFVRQH
      10      20      30      40      50      60

      70      80      90      100     110     120
m142.pep IDAAAVFRQDRNDSRTPVDAQHHGRRRLVGNRRDRRHCAVTPCRTVCRDDMNACRARCH
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
g142      IDAAAVFRQDRNDSRTPVYAQHHGRRRLVGNRRDRRHCAVTPCRTVCRDDMNACRTGCH
      70      80      90      100     110     120

      130     140     150     159
m142.pep RITERSLKIFLQIRHFSPLNCPPLYKNAAHKASPHVQQFX
      |||:|||||:|||||:|||||:|||||:|||||:|||||:
g142      RITERSLKSFLLQIRHFSPLNRPPLYKNAAHKASPHVQQFX
      130     140     150
```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 563>:

```
a142.seq
1  ATGCGTGCCG ATTTTCATGTT TGCCGACAAT ATGCCCCGTGC AGGTGCGCCA
51  ACGCGCCCTC TATTTCAAGT TGTCCCGTTT TGCCGCGATG CCAGATGTGG
101 TAGGCAAACC GCTCTTCGGG CGACAGGCCG GTCAGCCCGG CAAAATGTTC
151 GGCAACATCC TGATGTTCGT CCGCCAGCGT ATTGATGCAG AGGCTGCCGT
201 TTTCCGACAG GATCGGAATG ATTCGCGCAC TCCGGTTGAT GCACAGCATC
251 ACGGTCGGCG GCTCGTCCGT AACCGGCGCA ACCGCCGTCA TTGTAATGCC
301 GTAACGCCCT GCCGCACCGT CTGTCGTGAT GACATGAACG CCTGCCGCGC
351 AGGATGCCAT CGCATCACGG AACGAAGTTT GAAAAGTTT CTGCAAATCC
401 GCCATTTTTC CCCTTTAAAC TGTCCCTAT ATAAGAATGC TGCACACAAG
451 GCACCCCCCA TGTGCAGCAG TTCTGATTCA AAAAGCCGTC GGTCCGACAT
501 TTCCGCGCGT TACGGCGTAT TACGAGTTCA ACGCATCCTC GATTTTGGCA
551 AGTTCTGCCA ACAGGTCTTT AAGCAGCAGC ATTTTCTCGC GGCCAGCAC
601 TTCCTCGATA GCGTCGTAAC GCTCGTCCAC TTCTTCGCGG ATTTCTCAT
651 ACAGCTTCTC GCCCTCGGCA GTCAGCTTCA GAAAACACG TCGTTGGTCG
701 TTGGAAGGTT TCAGGCGGAC AACCAAACCC GCTTTTCAA GCGGGGTGAG
751 GATACCGGTC AGGCTGGGGC GCAAATGCA CGCCTGATTC GCCAAATCTT
```

801 GAAAGTCCAG CGTGCCGTTT TCCGCCAAAA GACGGATAAT CCGCCATTGC
 851 TGATCGGTAA TATTCGCCTG ATTCAGAATA GGCCTGAATT GGGTCATCAG
 901 GGCTTCCCTT GCCTGTATCA GACCGATATT GATAGACGCA TGTTTTGA

This corresponds to the amino acid sequence <SEQ ID 564; ORF 142.a>:

a142.pep
 1 MRADFMFADN MPVQVRQAL YFKLSRFAAM PDVVGKPLFG RQAGQPGKMF
 51 GNILMFVRQR IDAEAAVFRQ DRNDSRTPVD AQHHGRRLLVR NRRNRRHCNA
 101 VTPCRTVCRD DMNACRTGCH RITERSLKSF LQIRHFSPLN CPLYKNAAHK
 151 APPMCSSSDS KSRRSDISAR YGVLVRQIRL DFGKFCQQVF KQHFLLAAQH
 201 FLDSVVTLVH FFADEFLIQLL ALGSQLOKNT SLVVGRFQAD NQTRFFKAGQ
 251 DTGQAGAQNA RLIRQILKVQ RAVFRQKTDN PLLIGNIRL IQNRPELGHQ
 301 GFPCLYQTDI DRRMF*

m142/a142 96.1% identity in 153 aa overlap

	10	20	30	40	50	60
m142.pep	MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMF	GNILMFVRQR				
a142	MRADFMFADNMPVQVRQALYFKLSRFAAMPDVVGKPLFGRQAGQPGKMF	GNILMFVRQR				
	70	80	90	100	110	120
m142.pep	IDAEAAVFRQDRNDSRTPVDAQHHGRRLLVGNRRDRHCNAVTPCRTVCRDDMNACRARCH					
a142	IDAEAAVFRQDRNDSRTPVDAQHHGRRLLVGNRRDRHCNAVTPCRTVCRDDMNACRARCH					
	130	140	150	159		
m142.pep	RITERSLKIFLQIRHFSPLNCPPLYKNAAHKASPHVQQFX					
a142	RITERSLKIFLQIRHFSPLNCPPLYKNAAHKASPHVQQFX					
	130	140	150	160	170	180
a142	DFGKFCQQVFKQHFLLAAQHFLDSVVTLVHFFADEFLIQLLALGSQLOKNTSLVVGRFQAD					
	190	200	210	220	230	240

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 565>:

g143.seq
 1 ATGTTGAGCT TCGGCTATCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAAAG
 51 CTCGCAGATG AGCCGCAATT TTCAAACGCT AGGCGCAGAC CCGCACAAAT
 101 TGGGCTGGTT TTTCATCCTG CCGCCGCTGG CCGGGATGCT GGTTCAGCCG
 151 ATAGTGgGCT ACTACTCAGA CCGCACTTGG AAGCCGCGCT TGGGCGGCGC
 201 CCGCCTGCCG TATCTGCTTT ACGGCACGCT GATTGCGGTC ATCGTGATGA
 251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
 301 GCCTTGTCGT TCGGCGCGCT GATGATTGCG CTGTTGGACG TGTGCTCGAA
 351 TATGGCGATG CAGCCGTTTA AGATGATGGT CCGCGATATG GTCAACGAGG
 401 AGCAGAAAAG CTACGCCTAC GGGATTCAA GTTTCTTAGC GAATACGGAC
 451 GCGGTTGTGG CAGCGATTCT GCCGTTTGTG TTcgcgata TCGGTTTGGC
 501 GAACACTGCC GAGAAAGCGG TTGTGCCACA AACCGTGGTC GTAGCATTCT
 551 ATGTGGGTGC GCGGTACTG ATTATTACCA GTGCGTTCAC AATCTCCAAA
 601 GTCAAGAAT ACGACCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
 651 CGCCGCGAAT CAGGAAAAAG CCAACTGGTT CGAACTCTTA AAAACCGCGC
 701 CTAAAGTGTT TTGGACGGTT ACTCCGGTAC AGTTTTTCTG CTGGTTCGCC
 751 TTCCGGTATA TGTGGACTTA CTCGGCAGGC GCGATTGCAG AAAACGCTCG
 801 GCACACTACC GATGCGTCTT CCGTAGGCCA TCAGGAGGCG GGCAACCGGT
 851 ACGGCGTTTT GCGGCGGCTG TAGTCGGTTG CCGCGGTGAT TTGTTCTGTT
 901 ATTCTGGCAA AAGTACCGAA TAAATACCAT AAGGCGGGTT ATTTGCGCTG
 951 TTTGCGTTG GCGGCGCTCG GTTCTTCTC TATCTTCTC ATCTACAATC
 1001 AATACGCACT CATCCTGTCT TATATCTTAA TCGGCATCGC TTGGGCGGGC
 1051 ATTATCACTT ATCCGCTGAC GATTGTGGCC AACGCTTGT CGGGCAAAACA
 1101 CATGGATACT TATTGGGGCC TGttaacgg ctctgtCTGT ATGCcgcaaa
 1151 tcgTcgctTC GctgttgAGT TTCGTGCTTT TCCCGATGCT GGGCGGCCAT

405

1201 CAGGCAACCA TGTCTTGGT TGCAGGCGCA GTCTTGCTGC TGGGAGCCTT
 1251 CTCAGTCTGT CTGATTAAAG AGATCCACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 566; ORF 143.ng>:

g143.pep
 1 MLSFGYLGVO TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
 51 IVGYSDRTW KPRLGRRRLP YLLYGTLIIV IVMILMPNSG SFGFGYASLA
 101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKSYAY GIQSFLANTD
 151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL IITSFTISK
 201 VKEYDPETYA RYHGIDVAAN QEKANWFELL KTAPKVFETV TPVQFFCWFA
 251 FRYMWYTSAG AIAENVWHTT DASSVGHQEA GNRYGVLAHV *SVAVICSF
 301 ILAKVPNKYH KAGYFGCLAL GALGFFSIFV IYNQYALILS YILIGIAWAG
 351 IITYPLTIVA NALSGKHMDT YLGLFNQSVV MPQIVASLLS FVLFPMLGGH
 401 QATMFLVAGA VLLGAFSVC LIKEIHGGV*

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 567>:

m143.seq
 1 ATGCTCAGTT TCGGCTTTCT CGGCGTTCAG ACGGCCTTTA CCCTGCAAA
 51 CTGCAAAATG AGCCGCATTT TTCAAACGCT AGGCGCAGAC CCGCACAATT
 101 TGGGCTGGTT TTTTCATCCTG CCGCCGCTGG CGGGGATGCT GGTGCAGCCG
 151 ATTGTCGGCC ATTACTCCGA CCGCACTTGG AAGCCGCGTT TGGGCGGCCG
 201 CCGTCTGCCG TATCTGCTTT ATGGCACGCT GATTGCGGTT ATTGTGATGA
 251 TTTTGATGCC GAACTCGGGC AGCTTCGGTT TCGGCTATGC GTCGCTGGCG
 301 GCTTTGTCTG TCGGCGCGCT GATGATTGCG CTGTTAGACG TGTCGTCAAA
 351 TAGGCGCATG CAGCCGTTTA AGATGATGGT CGGCGACATG GTCACGAGG
 401 AGCAGAAAAG CTACGCCTAC GGGATTCAAA GTTTCTTAGC AAATACGGG
 451 GCGGTCGTGG CGGCGATTCT GCCGTTGTG TTTGCGTATA TCGGTTGGC
 501 GAACACCGCC GAGAAAGGCG TTGTGCCGCA GACCGTGGTC GTGGCGTTTT
 551 ATGTGGGTGC GCGGTTGCTG GTGATTACCA GCGCGTTCAC GATTTTCAAA
 601 GTGAAGGAAT ACGATCCGGA AACCTACGCC CGTTACCACG GCATCGATGT
 651 CGCCGCGAAT CAGGAAAAAG CCAACTGGAT CGAACTCTTG AAAACCGCGC
 701 CTAAGGCGTT TTGACGCTT ACTTTGGTGC AATTCTCTG CTGGTTCGCC
 751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
 801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAACTGGT
 851 ACGGCGTTTT GCGGCGGCTG CAGTCGTTG CCGCGGTGAT TTGTTCTGTT
 901 GTATTGCGCA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTGCGCTG
 951 TTTGCTTTG GCGCGCGCTG GCTTTTCTC CGTTTCTTC ATCGGCAACC
 1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
 1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
 1101 TATGGGCACT TACTTGGGCT TGTTTAACGG CTCTATCTGT ATGCCTCAAA
 1151 TCGTCCGCTT GCTGTTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
 1201 CAGGCCACTA TGTTCTTGGT AGGGGCGCTC GTCCTGCTGC TGGGCGCGTT
 1251 TTCCGTGTTT CTGATTAAAG AAACACACGG CGGGGTTTGA

This corresponds to the amino acid sequence <SEQ ID 568; ORF 143>:

m143.pep
 1 MLSFGYLGVO TAFTLQSSQM SRIFQTLGAD PHNLGWFFIL PPLAGMLVQP
 51 IVGHYSDRTW KPRLGRRRLP YLLYGTLIIV IVMILMPNSG SFGFGYASLA
 101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQKGYAY GIQSFLANTG
 151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSFTIFK
 201 VKEYDPETYA RYHGIDVAAN QEKANWIELL KTAPKAFETV TLVQFFCWFA
 251 FQYMWYTSAG AIAENVWHTT DASSVGYQEA GNWYGVLAHV QSVAVICSF
 301 VLAQVPNKYH KAGYFGCLAL GALGFFSVFF IGNOYALVLS YTLIGIAWAG
 351 IITYPLTIVT NALSGKHMDT YLGLFNQSVV MPQIVASLLS FVLFPMLGGL
 401 QATMFLVGGV VLLGAFSVC LIKEIHGGV*

Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from *N. gonorrhoeae*

m143 / g143 93.9% identity in 429 aa overlap

10 20 30 40 50 60

406

m143.pep	MLSFGFLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW
g143	MLSFGYLGVTQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGYYSDRTW
	10 20 30 40 50 60
m143.pep	70 80 90 100 110 120
	KPRLGGRRLPYLLYGTIAVIVMILPNSGSGFGYASLAALSFGALMIALLDVSSNMAM
g143	KPRLGGRRLPYLLYGTIAVIVMILPNSGSGFGYASLAALSFGALMIALLDVSSNMAM
	70 80 90 100 110 120
m143.pep	130 140 150 160 170 180
	QPFKMMVGDVNEEQKGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTAEGVVPQTVV
g143	QPFKMMVGDVNEEQKSYAYGIQSFLANTDAVVAAILPFVFAYIGLANTAEGVVPQTVV
	130 140 150 160 170 180
m143.pep	190 200 210 220 230 240
	VAFYVGAALLVITSAFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTA PKAFWTV
g143	VAFYVGAALLIITSAFTISKVKEYDPETYARYHGIDVAANQEKANWFELLKTA PKVFWTV
	190 200 210 220 230 240
m143.pep	250 260 270 280 290 300
	TLVQFFCWFQYMWYTSAGAIENVWHTTDASSVGYQEAGNRYGVLAQVSVAAVICSF
g143	TPVQFFCWFQYMWYTSAGAIENVWHTTDASSVGHQEAGNRYGVLAQVSVAAVICSF
	250 260 270 280 290 300
m143.pep	310 320 330 340 350 360
	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT
g143	ILAKVPNKYHKAGYFGCLALGALGFFSIFFIYNQYALILSYTLIGIAWAGIITYPLTIVA
	310 320 330 340 350 360
m143.pep	370 380 390 400 410 420
	NALSGKHMGTYLGLFNGSICMPQIVASLLSEVLFPMLGGLQATMFLVGGVLLLGAFSVF
g143	NALSGKHMDTYLGLFNGSVCMQIVASLLSEVLFPMLGGHQATMFLVAGAVLLLGAFSVC
	370 380 390 400 410 420
m143.pep	430
	LIKETHGGVX
g143	LIKEIHGGVX
	430

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 569>:

a143.seq

1	ATGCTCAGTT	TCGGCTTTCT	CGGCGTTCAG	ACGGCCTTTA	CCCTGCAAAG
51	CTCGCAGATG	AGCCGCATCT	TCCAGACGCT	CGGTGCCGAT	CCGCACAGCC
101	TCGGCTGGTT	CTTTATCCTG	CCGCCGCTGG	CGGGGATGCT	GGTGCAGCCG
151	ATTGTCGGCC	ATTACTCCGA	CCGCACTTGG	AAGCCGCGTT	TGGGCGGCCG
201	CCGTCGCGCC	TATCTGCTTT	ATGGCAGCCT	GATTGCGGTT	ATTGTGATGA
251	TTTTGATGCC	GAACGCGGC	AGCTTCGGTT	TCGGCTATGC	GTGCGTGCGC
301	GCTTTGTCGT	TCGGCGCGCT	GATGATTGCG	CTGTTAGACG	TGTCGTCAAA
351	TATGGCGATG	CAGCCGTTTA	AGATGATGGT	CGGCGACATG	GTCAACGAGG
401	AGCAGAAAGG	CTACGCCTAC	GGGATTCAAA	GTTTCTTAGC	GAATACGGGC
451	GCGGTCGTGG	CGGCGATTCT	GCCGTTTGTT	TTTGCGTATA	TCGGTTTGCC
501	GAACACCGCC	GAGAAAGGCG	TTGTGCCGCA	GACCGTGCTC	GTGCGGTTTT
551	ATGTGGGTGC	GCGGTGCTG	GTGATTACCA	GCGCGTTCAC	GATTTTCAAA
601	GTGAAGGAAT	ACAATCCGGA	AACCTACGCC	CGTTACCACG	GCATCGATGT
651	CGCCGCGAAT	CAGAAAAAG	CCAACGGAT	CGAACTCTTG	AAAACCGCGC
701	CTAAGGCGTT	TTGGACGGTT	ACTTTGGTGC	AATTCTTCTG	CTGGTTCGCC

```

751 TTCCAATATA TGTGGACTTA CTCGGCAGGC GCGATTGCGG AAAACGTCTG
801 GCACACCACC GATGCGTCTT CCGTAGGTTA TCAGGAGGCG GGTAAC TGGT
851 ACGGCGTTTT GCGGCGGGTG CAGTCGGTTG CGGCGGTGAT TTGTTTCGTTT
901 GTATTGGCGA AAGTGCCGAA TAAATACCAT AAGGCGGGTT ATTTTCGGCTG
951 TTTGGCTTTG GCGGCGCTCG GCTTTTCTC CGTTTTCTC ATCGGCAACC
1001 AATACGCGCT GGTGTTGTCT TATACCTTAA TCGGCATCGC TTGGGCGGGC
1051 ATTATCACTT ATCCGCTGAC GATTGTGACC AACGCCTTGT CGGGCAAGCA
1101 TATGGGCACT TACTTGGGCC TGTTAACGG CTCTATCTGT ATGCCGCAAA
1151 TCGTCGCTTC GCTGTGAGT TTCGTGCTTT TCCCTATGCT GGGCGGCTTG
1201 CAGGCCACTA TGTTCTTGGT AGGGGGCGTC GTCTGCTGC TGGGCGCGTT
1251 TTCCGTGTTT CTGATTAAAG AAACACACGG CGGGGTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 570; ORF 143.a>:

```

a143.pep
1  MLSFGLGVQ TAFTLQSSQM SRIFQTLGAD PHSLGWFFIL PPLAGMLVQP
51  IVGHYSDRTW KPRLGRRLP YLLYGTLIIV IVMILMPNSG SFGFGYASLA
101 ALSFGALMIA LLDVSSNMAM QPFKMMVGDM VNEEQGYAY GIQSFLANTG
151 AVVAAILPFV FAYIGLANTA EKGVPQTVV VAFYVGAALL VITSAFTIFK
201 VKEYNPETYA RYHGIDVAAN QEKANWIELL KTAPKAFWTV TLVQFFCWFA
251 FQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAAVICSF
301 VLAKVPNKYH KAGYFGCLAL GALGFFSVFF IGNQYALVLS YTLIGIAWAG
351 IITYPLTIVT NALSGKHMGT YLGLFNGSIC MPQIVASLLS FVLFPMLGGL
401 QATMFLVGGV VLLGAFSVF LIKETHGGV*

```

m143/a143 99.5% identity in 429 aa overlap

	10	20	30	40	50	60
m143.pep	MLSFGLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW					
a143	MLSFGLGVQTAFTLQSSQMSRIFQTLGADPHNLGWFFILPPLAGMLVQPIVGHYSDRTW					
	10	20	30	40	50	60
m143.pep	KPRLGRRLPYLLYGTLIIVIMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
a143	KPRLGRRLPYLLYGTLIIVIMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
	70	80	90	100	110	120
m143.pep	KPRLGRRLPYLLYGTLIIVIMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
a143	KPRLGRRLPYLLYGTLIIVIMILMPNSGSFGFGYASLAALSFGALMIALLDVSSNMAM					
	70	80	90	100	110	120
m143.pep	QPFKMMVGDMVNEEQGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTA EKGVPQTVV					
a143	QPFKMMVGDMVNEEQGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTA EKGVPQTVV					
	130	140	150	160	170	180
m143.pep	QPFKMMVGDMVNEEQGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTA EKGVPQTVV					
a143	QPFKMMVGDMVNEEQGYAYGIQSFLANTGAVVAAILPFVFAYIGLANTA EKGVPQTVV					
	130	140	150	160	170	180
m143.pep	VAFYVGAALLVITSAFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
a143	VAFYVGAALLVITSAFTIFKVKEYNPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
	190	200	210	220	230	240
m143.pep	VAFYVGAALLVITSAFTIFKVKEYDPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
a143	VAFYVGAALLVITSAFTIFKVKEYNPETYARYHGIDVAANQEKANWIELLKTAPKAFWTV					
	190	200	210	220	230	240
m143.pep	TLVQFFCWFAFQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAAVICSF					
a143	TLVQFFCWFAFQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAAVICSF					
	250	260	270	280	290	300
m143.pep	TLVQFFCWFAFQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAAVICSF					
a143	TLVQFFCWFAFQYMWYSAG AIAENVWHTT DASSVGYQEA GNWYGVLA AVQSVAAVICSF					
	250	260	270	280	290	300
m143.pep	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
a143	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
	310	320	330	340	350	360
m143.pep	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
a143	VLAKVPNKYHKAGYFGCLALGALGFFSVFFIGNQYALVLSYTLIGIAWAGIITYPLTIVT					
	310	320	330	340	350	360
m143.pep	NALSGKHMGT YLGLFNGSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLGAFSVF					
a143	NALSGKHMGT YLGLFNGSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLGAFSVF					
	370	380	390	400	410	420
m143.pep	NALSGKHMGT YLGLFNGSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLGAFSVF					
a143	NALSGKHMGT YLGLFNGSICMPQIVASLLSFVLFPMLGGLQATMFLVGGVVLLGAFSVF					

408

```

a143      |||||
          NALSGKHMGTYLGLFNGSICMPQIVASLLSFVLPMLGGLQATMFLVGGVVLLLGAFSVF
          370      380      390      400      410      420

          430
m143.pep  LIKETHGGVX
          |||||
a143      LIKETHGGVX
          430

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 571>:

```

g144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGGGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGTGC GTCTTCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGGCGTG
151 CGCGAAAACC CCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGTAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CGGTTGCAC GGGCTGGCCG TTACCcgtTT
351 CAACGCGGTG GCGGCAGACG GccgacggTt atCCCAACGA TTTGgatatT
401 TCctaccgCT TGGACGAGGA CGGCCGGCTT ACCGTtacct ATCGCGCCAC
451 CGCgctCGGC GACACGGTGT TCGACCGGAC GCTGCACATT TACTGGCGGC
501 TGGACGCGGG CCTGCACGAT GCGGTTCTGC ATATTCCGCA GGGCGGACAT
551 ATTCCGGCCG ATGCCGAAAA ACTGCCCGTC TTAACGGTTT CAGACGGCCT
601 CGAAGTATTT GA

```

This corresponds to the amino acid sequence <SEQ ID 572; ORF 144.ng>:

```

g144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTC VFVLDLGGIV QEFVSLADGV
51  RENPVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRRLSQR FGYFLPLGRG RPAYRYLSRH
151 RARRHGVPRD AAHLLAAGRG PARCGSAYSA GRTYSGRCRK TARLNGFRRP
201 RSI*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 573>:

```

m144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGTCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGGCGTG
151 CGCGAAAACC TCGTGGTGTC GTTCGATGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAAC AGATAGGGCG CGTGGCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CGGTTGCAC GGGCTGGCCG TTACCGGTTT
351 CAACGCGGTG GCGGCAGACG GCCGTTGCGT GGTGCTGCGC AGCCGCCTGg
401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTGg ATATTTCCTA
451 CCGCTTGGAC GAGGACGACC GGCTTACCGT TAaCTATCGC GCCACCGCGC
501 TCGGCGACAC GGTGTTTCGAC CCGACGCTGC ACATTTACTG GCGGCTGGAC
551 GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATGCC
601 GGCCGATGCC GAAAACTGC CCGTCTCAAC GGTTCAGAC GACCTCGAAG
651 TATTGTA

```

This corresponds to the amino acid sequence <SEQ ID 574; ORF 144>:

```

m144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFVSLADGV
51  RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLATVGRR L SQRFGFGYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAAHLL AAGRGPARGC SAYSAGRTYA
201 GRCRK TARLN GFRRPRSI*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m144 / g144 91.3% identity in 218 aa overlap

	10	20	30	40	50	60
m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD					
g144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTCVFVLDLGGIVQEFSVLADGVRENPFVVSFDD					
	10	20	30	40	50	60
	70	80	90	100	110	120
m144.pep	AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV					
g144	AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV					
	70	80	90	100	110	120
	130	140	150	160	170	180
m144.pep	AADGRSVVLRSLATVGRRLSQRFGFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAHLL					
g144	AAD-----GRRLSQRFG--YFLPLGRGRPAYRYLSRHRARRHGVRPDAHLL					
		130	140	150	160	
	190	200	210	219		
m144.pep	AAGRGPARGSAYSAGRTYAGRCRKRTARLNGFRRPRSIX					
g144	AAGRGPARGSAYSAGRTYSGRCRKRTARLNGFRRPRSIX					
	170	180	190	200		

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 575>:

```

a144.seq
1  ATGAGCGATA CCCCCGCTAC CCGCGATTTC GGCCTGATCG ACGGGCGTGC
51  CGTAACCGGC TATGTGCTGT CCAACCGGCG TGGTACGCGT GTCTGCGTGC
101 TGGACTTGGG CGGGATTGTG CAGGAATTTT CCGTTTGGC AGACGGCGTG
151 CGCGAAAACC TCGTGGTGTC GTTCGACGAT GCGGCTTCCT ATGCGGACAA
201 TCCGTTTCAG ATTAACAAGC AGATAGGGCG CGTGCCCGGA CGCATCCGCG
251 GTGCGGCGTT CGACATCAAC GGCAGGACTT ACCGCGTGGA GGCCAACGAA
301 GGCAGGAACG CGCTGCACGG CGGTTTCGAC GGGCTGGCCG TTACCCGTTT
351 CAACGCGGTT GCGGCAGACG GCCGTTCCGT GGTGCTGCGC AGCCGCCTG.
401 CAACAGTCGG CCGACGGTTA TCCCAACGAT TTGGATTGG ATATTTCCTA
451 CCGCTTGGAC GAGGACGACC GGCTTACCGT TACCTATCGC GCCACCGCGC
501 TCGGCGACAC GGTGTTTCGAC CCGACGCTGC ACATTACTG GCGGCTGGAC
551 GCGGGCCTGC ACGATGCGGT TCTGCATATT CCGCAGGGCG GACATATTCC
601 GGCCGATGCC GAAAACTGC CCGTCTCAAC GGTTTCAGAC GACCTCGAAG
651 TATTGA

```

This corresponds to the amino acid sequence <SEQ ID 576; ORF 144.a>:

```

a144.pep
1  MSDTPATRDF GLIDGRAVTG YVLSNRRGTR VCVLDLGGIV QEFSVLADGV
51  RENLVVSFDD AASYADNPFQ INKQIGRVAG RIRGAAFDIN GRTYRVEANE
101 GRNALHGGSH GLAVTRFNAV AADGRSVVLR SRLXTVGRRL SQRFGFGYFL
151 PLGRGRPAYR YLSRHRARRH GVRPDAHLL AAGRGPARGC SAYSAGRTYS
201 GRCRKRTARLN GFRRPRSIX*

```

m144/a144 99.1% identity in 218 aa overlap

	10	20	30	40	50	60
m144.pep	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD					
a144	MSDTPATRDFGLIDGRAVTGYVLSNRRGTRVCVLDLGGIVQEFSVLADGVRENLVVSFDD					
	10	20	30	40	50	60
	70	80	90	100	110	120
m144.pep	AASYADNPFQINKQIGRVAGRIRGAADFINGRTYRVEANEGRNALHGGSHGLAVTRFNAV					

410

```

|||||
a144      AASYADNPFQINKQIGRVAGIRGAAFDINGRTYRVEANEGRNALHGGSHGLAVTRFNAV
              70          80          90          100          110          120

              130          140          150          160          170          180
m144.pep  AADGRSVVLRSLATVGRRLSQRFQFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
              |||||||
a144      AADGRSVVLRSLXTVGRRLSQRFQFGYFLPLGRGRPAYRYLSRHRARRHGVRPDAAHLL
              130          140          150          160          170          180

              190          200          210          219
m144.pep  AAGRGPARGCSAYSAGRTYAGRCRKRTARLNGFRRPRSIX
              |||||||
a144      AAGRGPARGCSAYSAGRTYSGRCRKRTARLNGFRRPRSIX
              190          200          210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 577>:

```

g146.seq
1  ATGAAGCAAA TCCCCCTCCG CCTTCTCCAG GTCGTCATTG ACCACGACAA
51  AGTCGAACAA TACGGAAGTGT TCGATTTCAT GCCTTGCCCTT CGACAGCCTC
101 CTTTGGATAA ctTCCCGACT GTCCGTCCCG CGCctTTGA GCGCGCGGCG
151 AAGCAGTCG AAAGAAGGCG GCAGGATAAA GATACCGACA GCTTCCGGCA
201 GCGCGTTCG AACCTGCGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAGCCTG CCGCCGCCAA CGCATTACG CCTCCGTGC TTGTGCCGTA
301 ATAGTTGCCG AATACGCTCTG CGTATTCCAA AAAAGCCTCC TCGCGGATAA
351 GCGATTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGGCG CGTCGTATGC GACACGGAAA CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TCGATAAGC GGACATATT
551 TTTACCTGTA TATTTTCCAA CCGATTGTAT CACAACGGAC ACCCTATTTC
601 ATATTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 578; ORF 146.ng>:

```

g146.pep
1  MKQIPLRLQ VVIDHDKVEQ YGLFDFMPCL RQPPLDNFPT VRPAPFEARG
51  KHVERRRQDK DTDSEFRQVA NLRRALNVDF QNHVIACRRQ RIHALRACAV
101 IVAEYVCVFQ KSLLRDKRFK LFFGNKVIMY AVCFAFTRRA RMRHGNAGT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQRTPYF
201 IFADAHILPL LF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 579>:

```

m146.seq
1  ATGGCGCAAA TCCTCCTCCG CTCGCGCCAA GTCGTCATTG ACCACGACAA
51  AGTCAAACAA TACGGAAGTGT TCGATTTCAT GCCTTGCCCTT CGACAGCCTC
101 CTTTGGATAA CTCCCGACT GTCCGTCCCG CGTCCGTGTA GCGCGCGGCG
151 AAGTACGTCG AAAGAAGGCG GCAGGATAAA GATGCCGACG GCTTCGGGCA
201 GCGCGTTCG AACCTGCGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAGCCTG CCGCCGCCAA CGCATTACA CCTCCGCGC CTGTGCCGTA
301 ATAGTTGCCA AATACGTCGG CGTATTCCAA AAAAGCCTCC TCGCGGATAA
351 GCGACTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGGCG CGTCGTGTGC GACACGGAAA CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TCGATAAGC GGACATATT
551 TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAATGGAC ACCCAGTTTC
601 CTATTGCCG ATGCCCATAT TTTGCCGCTA TTGTTTTGA

```

This corresponds to the amino acid sequence <SEQ ID 580; ORF 146>:

```

m146.pep
1  MAQILLRSRQ VVIDHDKVKQ YGLLDFMPCL RQPPLDNFPT VRPASVEARG
51  KYVERRRQDK DADGFGQOVA NLRRALNVDF QNHVIACRRQ RIHTLRACAV
101 IVAKYVGVFQ KSFLRDKRLK LFFGNKVIMY AVCFAFTRRA RRVRHGNAQT
151 VMVCQQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQWTPSF
201 LFADAHILPL LF*

```


Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m146 / g146 90.1% identity in 212 aa overlap

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVV	IDHDKVKYGLLD	FMPCLRQPPLD	NFPTVRPASV	EARGKYVERRR	QDK
g146	MKQIPLRLQV	VIDHDKVEQYGL	DFMPCLRQPPL	DNFPTVRPAP	FEARGKHVERR	RQDK
	10	20	30	40	50	60
m146.pep	DADGFGQRVAN	LRRALNVDFQNH	VIACRRQRIHT	LRACAVIVAKY	VGVFQKSLRD	KRLK
	:					
g146	DTDSFRQRVAN	LRRALNVDFQNH	VIACRRQRIHAL	RACAVIVAEYV	CVFQKSLLRD	KRFK
	70	80	90	100	110	120
m146.pep	LFFGNKVIMYAV	CFATRRARRVR	HGNAQTMVCQ	QPRHQGFARAG	SGRNDKDVA	FSIS
g146	LFFGNKVIMYAV	CFATRRARRMR	HGNAQTMVCQ	QPRHQGFARAG	SGRNDKDVA	FSIS
	130	140	150	160	170	180
m146.pep	GHIFYLYIFQPI	VSQWTPSFLFAD	AHILPLLEFX			
g146	GHIFYLYIFQPI	VSQRTPFIFAD	AHILPLLEFX			
	190	200	210			

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 581>:

```

a146.seq
1  ATGGCGCAAA TCCTCCTCCG CCCGCGCCAA GTCATCATTG ACCACGACAA
51  AATCGAACAA TACGGACTGT TCGATTTCAT GCCTTGCTT CGACAGCCTC
101 CTTTGGATAA CTTCCTGACT GTCCGTCCCG CGTCCGTTGA GACGCGCAGC
151 AAGCACATCG AAAGACGGCG GCAGGATAAA GATGCCGACG GCTTCGGGCA
201 GCGCATCTCG AACCTGAGCC GCGCCCTGAA CGTCGATTTC CAAAATCACG
251 TCATAACCTG CCGCCGCCAA CGCATTACA CCCTCCGCGC TTGTGCCGTA
301 ATAGTTGCCG AACACGTCCG CGTATTCCAA AAAAGCCTCC TGCGCGATAA
351 GCGACTCAAA CTCTTCTTTG GAAACAAAGT GATAATGTAC GCCGTTTGCT
401 TCGCCTTCAC GCGGCGGACG CGTCGTGTGC GACACGGAAA CGCGCAAACC
451 GTTATGGTTT GCCAACAGCC GCGACACCAG CGTGGTTTTC CCCGTGCCGG
501 AAGCGGCCGA AATGATAAAG ATGTTGCCTT TTCGATAAGC GGACATATTT
551 TTTACCTGTA TATTTTCCAG CCGATTGTAT CACAACGGAC ACCCGGTTTC
601 CTATTTGCCG ATGCCCATAT TTGCGCGCTA TTGTTTGA
  
```

This corresponds to the amino acid sequence <SEQ ID 582; ORF 146.a>:

```

a146.pep
1  MAQILLRPRO VIIDHKIEQ YGLDFMPCL RQPPLDNFPT VRPASVETRS
51  KHIERRQDK DADGFGQRI NLSRALNVDF QNHVITCRRQ RIHTLRACAV
101 IVAEHVRVFQ KSLLRDKRLK LFFGNKVIMY AVCFAFTRRT RVRHGNNAQT
151 VMVCQPRHQ RGFARAGSGR NDKDVAFSIS GHIFYLYIFQ PIVSQRTPGF
201 LFADAHILPL LF*
  
```

m146/a146 90.6% identity in 212 aa overlap

	10	20	30	40	50	60
m146.pep	MAQILLRSRQVV	IDHDKVKYGLLD	FMPCLRQPPLD	NFPTVRPASV	EARGKYVERRR	QDK
a146	MAQILLRPRQVI	IDHDKIEQYGL	DFMPCLRQPPL	DNFPTVRPAS	VETRSKHIERR	QDK
	10	20	30	40	50	60
	70	80	90	100	110	120

412

```

m146.pep  DADGFGQRVANLRRALNVDFQNHVIACRRQRIHTLRACAVIVAKYVGVFQKSFRLDKRLK
          |||||::|| |||||::|||:|||||::|||:|||||::|||:|||||
a146      DADGFGQRISNLSRALNVDFQNHVITCRRQRIHTLRACAVIVAEHVRVFQKSLRLDKRLK
          70      80      90      100     110     120

          130     140     150     160     170     180
m146.pep  LFFGNKVIMYAVCFATRARRVRHGNATVMVCQPRHQRGFARAGSGRNDKDVAFSIS
          |||||::|||:|||||::|||:|||||::|||:|||||::|||:|||||
a146      LFFGNKVIMYAVCFATRARRVRHGNATVMVCQPRHQRGFARAGSGRNDKDVAFSIS
          130     140     150     160     170     180

          190     200     210
m146.pep  GHIFYLYIFQPIVSQWTPSFLEDAHILPLLF
          |||||::|||:|||||::|||:|||||::|||:|||||
a146      GHIFYLYIFQPIVSQRTPGFLEDAHILPLLF
          190     200     210

```

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 583>:

```

g147.seq (partial)
1  ..ATGCGACGAG AAGCCAAAAT GGCACAAATC ACACTCAAAC CCATTGTTTT
51  ATCAATTCTT TTAATCAACA CACCCCTCCT CGCCCAAGCG CATGAAACTG
101 AGCAATCGGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCCG
151 CGCGCGACTT CGGGCTGCT GCACACTTCG ACCGCCTCCG ACAAATCAT
201 CTCCGGCGAT ACTTTGCGCC AAAAGCCGT CAACTTGGGC GACGCTTGG
251 ACGGCGTACC GGGCATCCAC GCTTCGCAAT ACGGCGCGG CGCATCCGCT
301 CCCGTTATTC GCGGTCAAAC GGGCAGACGG ATTAAAGTAT TGAACCATCA
351 CGGCGAAACG GGCATATGG CGGACTTTC TCCGATCAC GCCATTATGG
401 TAGATACCGC CTTGTCGCAA CAGGTTGAAA TCCTGCGCGG GCCGTTACG
451 CTCTGTACA GCTCGGgcaa tgtggccgGG GCTGGtcaat gttgccgatg
501 gAAAAAtccc ccaaaaAAtg cc..

```

This corresponds to the amino acid sequence <SEQ ID 584; ORF 147.ng>:

```

g147.pep (partial)
1  ..MRREAKMAQI TLKPIVLSIL LINTPLLAQA HETEQSVGLE TVSVVVKSRP
51  RATSGLLHTS TASDKIISGD TLRQKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIRGQTGRR IKVLNHHGET GDMADFSPDH AIMVDTALSQ QVEILRGPVT
151 LLYSSGNVAG AGQCCRWKNP PKNA..

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 585>:

```

m147.seq (partial)
1  ..CCGCATAAAA CTGAGCAATC GGTGGATTG GAAACGGTCA GCGTCGTCGG
51  CAAAAGCCGT CCGCGCGCCA CGTCGGGGCT GTTGCACT TCGACCGCCT
101 CCGACAAAAT CATCTCCGGC GATACCTGC GCCAAAAGC CGTCAACTTG
151 GGCACGCTT TAGACGGCGT ACCGGGCATC CACGCTTCGC AATACGGCGG
201 CGGCGCTCT GTCCTCGTCA TTCGCGTCA AACAGGCAGG CGGATTAAAG
251 TGTGGAACCA TCACGGCGAA ACAGGCGATA TGGCGGATT TTCGCCCGAT
301 CACGCCATTA TGGTAGATAC CGCCTTGTCG CAACAGGTCG AAATCCTGCG
351 CGGCGCGGT ACGCTCTTGT ACAGCTCGGG CAATGTGGCG GGGCTGGTCG
401 ATGTTGCCGA TGGCAAAATC CCCGAAAAA TGCCTGAAAA CGGCGTATCG
451 GGCGAACTCG GATTGCGTT GAGCAGCGGC AATCTGAAA AACTCACGTC
501 CGGCGGCATC AATATCGGTT TGGGCAAAA CTTTGTATTG CACACGGAAG
551 GGCTGTACCG CAAATCGGGG GATTACGCC TACCGCGTTA CCGCAATCTG
601 AAACGCCTGC CCGACAGCCA CGCCGATTG CAAACGGGCA GCATCGGGCT
651 GTCTTGGGTT GCGGAAAAAG GTTTTATCGG CGTAGCGTAC AGCGACGTC
701 GCGACCAATA TGGTCTGCCT GCCCAGAGC ACGAATACGA TGATTGCCAC
751 GCCGACATCA TCTGGCAAAA GAGCTTGATT AACAAACGCT ATTTACAGCT
801 TTATCCGCAC CTGTTGACCG AAGAAGACAT CGATTACGAC AATCCGGGCT
851 TGAGCTGCGG CTTCCACGAC GACGATAATG CACACGCACA CACCCACGC
901 GGCAGACCGT GGATAGACCT GCGCAACAAA CGCTACGAAC TCCGTGCCGA
951 ATGAAGCAA CCGTCCCG GTTTTGAAGC CCTGCGCGTA CACCTGAACC
1001 GCAACGACTA CCGCCACGAC GAAAAAGCAG GCGATGCAGT CGAAAACCTT
1051 TTTAACAAAC AAACGCAAAA CGCCCGCATC GAGTTGCGCC ACCAACCCAT

```

```

1101 AGGTCGTCTG AAAGGCAGCT GGGGCGTGCA ATATTTACAA CAAAAATCCA
1151 GTGCTTTATC TGCCATATCC GAAGCGGTTA AACAAACCGAT GCTGCTTGAC
1201 AACAAAGTGC AACATTACAG CTTTTTCGGT GTAGAACAGG CAAACTGGGA
1251 CAACTTCACG CTTGAAGGAG GCGTACGCGT GGAAAAACAA AAAGCCTCCA
1301 TTCAGTACGA CAAAGCATTG ATTGATCGGG AAAACTACTA CAACCAACCC
1351 CTGCCCCGACC TCGGCGCGCA CCGCCAAACC GCCCGCTCAT TCGCACTTTC
1401 GGGCAACTGG TATTTACGCG CACAACACAA ACTCAGCCTG ACCGCCTCCC
1451 ATCAGGAACG CCTGCCGTCA ACGCAAGAGC TGTACGCACA CGGCAAAACAC
1501 GTCGCCACCA ACACCTTTGA AGTCGGCAAC AAACACCTCA ACAAGAGCG
1551 TTCCAACAAT ATCGAACTCG CGCTGGGCTA CGAAGGCGAC CGCTGGCAAT
1601 ACAATCTGGC ACTCTACCGC AACCCTTCG GTAACTACAT TTACGCCAA
1651 ACCTTAAACG ACTGACGCGG CCCCAAATCC ATCGAAGACG ACAGCGAAAT
1701 GAAGCTCGTG CGCTACAACC AATCCGGCGC CGACTTCTAC GCGCGGGAAG
1751 GCGAAATCTA CTTCAAACCG ACACCGCGCT ACCGCATCGG CGTTTCCGGC
1801 GACTATGTAC GAGGCGCTCT GAAAAACCTG CCTTCCCTAC CCGGCAGAGA
1851 AGATGCCTAC GGCAACCGTC CTTTCATCGC ACAGGACGAC CAAAATGCC
1901 CCCGTGTTCC GGCTGCGCGC CTCGGCTTCC ACCTGAAAGC CTCGCTGACC
1951 GACCGTATCG ATGCCAATTT GGACTACTAC CGCGTGTTCG CCCAAACAA
2001 ACTCGCCCGC TACGAAACGC GCACGCCCGG ACACCATATG CTCAACCTCG
2051 GCGCAAACCTA CCGCCGCAAT ACGCGCTATG GCGAGTGGAA TTGGTACGTC
2101 AAAGCCGACA ACCTGCTCAA CCAATCCGTT TACGCCACCA GCAGCTTCT
2151 CTCTGATACG CCGCAAATGG GCGCAGCTT TACCGCGCGC GTGAACGTGA
2201 AGTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 586; ORF 147>:

```

m147.pep (partial)
1 ..PHKTEQSVDL ETVSVVGKSR PRATSGLLHT STASDKIISG DTLRQKAVNL
51 GDALDGVPGI HASQYGGGAS APVIRGQTGR RIKVLNHHGE TGDMAFSPD
101 HAIMVDTALS QQVEILRGPV TLLYSSGNVA GLVDVADGKI PEKMPENGVS
151 GELGLRLSSG NLEKLTSGGI NIGLGKNEVL HTEGLYRKSG DYAVPRYRNL
201 KRLPDSHADS QTSIGLSWV GEKGFIVGAY SDRRDQYGLP AHSHEYDDCH
251 ADIIWQKSLI NKRYLQLYPH LLTEEDIDYD NPGLSCGFHD DDNAHAHTHS
301 GRPWIDLRNK RYELRAEWKQ PFPGFALRV HLNRRNDYRD EKAGDAVENF
351 FNNQTONARI ELRHQPIGRL KGSWGVQYLQ QKSSALSAIS EAVKQPMLLD
401 NKVQHYSFFG VEQANWDNFT LEGGVRVEKQ KASIQYDKAL IDRENYYNHP
451 LPDLGAHRQT ARSFALSGNW YFTPQHKLSL TASHQERLPS TQELYAHGKH
501 VATNTFEVGN KHLNERSNN IELALGYEGD RWQYNLALYR NRGNYIYQAQ
551 TLNDGRGPKS IEDDSEMKLV RYNQSGADFY GAEGEYFKP TPRYRIGVSG
601 DYVRGRLLKNL PSLPGREDAY GNRPFIAQDD QNAPRVPAAR LGFHLKASLT
651 DRIDANLDYY RVFAQNKLAR YETRTPGHHM LNLGANYRRN TRYGEWNWYV
701 KADNLLNQSV YAHSSFLSDT PQMGRSFTGG VNVKF*

```

Computer analysis of this amino acid sequence gave the following results:
Homology with a predicted ORF from *N. gonorrhoeae*

m147 / g147 92.3% identity in 142 aa overlap

```

m147.pep          10      20      30
                  PHKTEQSVDL ETVSVVGKSR PRATSGLLHTS
g147              MRREAKMAQITLKPIVLSILLINTPLLAQAHETEQSVGLETSVSVVGKSR PRATSGLLHTS
                  10      20      30      40      50      60
m147.pep          40      50      60      70      80      90
TASDKIISGDTLRQKAVNLGDALDGVPGI HASQYGGGASAPVIRGQTGRRIKVLNHHGET
g147              TASDKIISGDTLRQKAVNLGDALDGVPGI HASQYGGGASAPVIRGQTGRRIKVLNHHGET
                  70      80      90      100     110     120
m147.pep          100     110     120     130     140     150
GDMADFSPDHAIMVDTALSQQVEILRGPV TLLYSSGNVAGLVDVADGKIPEKMPENGVS
g147              GDMADFSPDHAIMVDTALSQQVEILRGPV TLLYSSGNVAGAGQCCRWKNPPKNA

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 587>:

```
a147.seq
1  ATGCGACGAG AAGCCAAAAT GGCACAAACT ACACCTCAAAC CCATTGTTTT
51  ATCAATCTTT TTAATCAACA CACCCTCCTT CTCCCAAGCG CATGGAACCTG
101 AGCAATCAGT GGGCTTGGAA ACGGTCAGCG TCGTCGGCAA AAGCCGTCGG
151 CGCGCCACTT CGGGGCTGCT GCACACTTCT ACCGCTCCG ACAAAATCAT
201 CAGCGGCGAC ACCTTGCAGC AAAAAGCCGT CAACTTGGGT GATGCTTTAG
251 ACGGCGTACC GGGCATTCAT GCCTCGCAAT ACGGCGGCGG CGCATCCGCT
301 CCCGTTATTC GCGGTCAAAC AGGCAGACGG ATTAAGTGT TGAACCATCA
351 CGGCGAAACG GGCACATGG CGGACTTCTC TCCAGACCAT GCAATCATGG
401 TGGCAGCGC CTTGTCGCAA CAGGTCGAAA TCCTGCGCGG TCCGGTTACG
451 CTCTTGATCA GCTCGGGCAA TGTGGCGGGG CTGGTCGATG TTGCGATGG
501 CAAAATCCCC GAAAAAATGC CTGAAAACGG CGTATCGGGC GAACTCGGAT
551 TCGGTTTGAG CAGCGGCAAT CTGAAAAAAC TCACGTCCGG CGGCATCAAT
601 ATCGGTTTGG GCAAAACTTT TGTATTGCAC ACGGAAGGGC TGTACCGCAA
651 ATCGGGGGAT TACCGGTAC CGCGTTACCG CAATCTGAAA CGCTGCCCCG
701 ACAGCCACGC CGATTGCGAA ACGGGCAGCA TCGGGCTGTC TTGGGTTGGC
751 GAAAAAGGCT TTATCGGCGC AGCATACAGC GACGTCGCG ACCAATATGG
801 TCTGCCTGCC CACAGCCACG AATACGATGA TTGCCACGCC GACATCATCT
851 GCAAAAAGAG TTTGATTAA ACACGCTATT TGCACTTTA TCCGCACCTG
901 TTGACCGAAG AAGACATCGA TTACGACAA CCGGGCTTGA GCTGCGGCTT
951 TCACGACGAC GATGATGCAC ACGCCCATGC CCACAACGGC AAACCTTGGA
1001 TAGACCTGCG CAACAAACGC TACGAACTCC GCGCGAATG GAAGCAACCG
1051 TTCCCGGTTT TTGAAGCCCT GCGGTACAC CTGAACCGCA ACGACTACCG
1101 CCACGACGAA AAAGCAGGCG ATGCAGTAGA AAACCTTTT AACAACCAA
1151 CGCAAAACGC CCGTATCGAG TTGCGCCACC AACCCATAGG CCGTCTGAAA
1201 GGCAGCTGGG GCGTCAATA TTTGGGACAA AAATCCAGTG CTTATCTGTC
1251 CACATCCGAA GCGGTCAAAC AACCGATGCT GCTTGACAA AAAGTGCAAC
1301 ATTACAGCTT TTTGCGTGTG GAACAGGCAA ACTGGGACAA CTTACGCTT
1351 GAAGGCGGCG TACGCGTGGA AAAACAAAA GCCTCCATCC GCTACGACAA
1401 AGCATTGATT GATCGGGAAA ACTACTACAA CCATCCCCTG CCCGACCTCG
1451 GCGCGCACCG CCAAAACGCC CGCTCATTCG CACTTTCGGG CAACTGGTAT
1501 TTTACGCGAC AACACAAACT CAGCCTGACC GCCTCCCATC AGGAACGCCT
1551 GCCGTCAACG CAAGAGCTGT ACGCACACGG CAAACACGTC GCCACCAACA
1601 CCTTTGAAGT CGGCAACAAA CACCTCAACA AAGAGCGTTC CAACAATATC
1651 GAACTCGCGC TGGGCTACGA AGGCGACCGC TGGCAATACA ATCTGGCACT
1701 CTACCGCAAC CGCTTCGGCA ACTACATTTA CGCCCAACC TTAACGACG
1751 GACGCGGCCC CAAATCCATC GAAGACGACA GCGAAATGAA GCTCGTGC
1801 TACAACCAAT CCGGTGCGGA CTTCTACGGC GCGGAAGGCG AAATCTACTT
1851 CAAACCGACA CCGCGCTACC GCATCGGCGT TTCCGGCGAC TATGTACGAG
1901 GCGCTCTGAA AAACCTGCCT TCCCTACCG GCAGGGAAGA CGCCTACGGC
1951 AACCGCCAC TCATTGCCCA AGCCGACCAA AACGCCCTC GCGTTCCGGC
2001 TGCGCGCCTC GCGTCCACC TGAAAGCCTC GCTGACCGAC CGCATCGATG
2051 CCAATTTGGA CTACTACGCG GTGTTGCGCC AAAACAACT CGCCCGCTAC
2101 GAAACGCGCA CGCCGGACA CCATATGCTC AACCTCGGCG CAAACTACCG
2151 CCGCAATACG CGCTATGGCG AGTGGAATTG GTACGTCAAA GCCGACAACC
2201 TGCTCAACCA ATCCGTTTAC GCCACAGCA GCTTCCTCTC TGATACGCCG
2251 CAAATGGGCC GCAGCTTTAC CGCGGCGGTG AACGTGAAGT TTAA
```

This corresponds to the amino acid sequence <SEQ ID 588; ORF 147.a>:

```
a147.pep
1  MRREAKMAQT TLKPVLISIL LINTPLLSQA HGTEQSVGLE TVSVVGKSRP
51  RATSGLLHTS TASDKIISGD TLROKAVNLG DALDGVPGIH ASQYGGGASA
101 PVIRGQTGRR IKVLNHHGET GDMADFS PDH AIMVDSALSQ QVEILRGPVT
151 LLYSSGNVAG LVDVADGKIP EKPENGVSQ ELGLRLSSG LEKLTSGGIN
201 IGLGNFVLH TEGLYRKSGD YAVPRYRNLK RLPDSDHSDQ TGSIGLSWVG
251 EKGFIGAAYS DRRDQYGLPA HSHEYDDCHA DIIWQKSLIN KRYLQLYPHL
301 LTEDIDYDN PGLSCGFHDD DDAHAHAHNG KPWIDLNRKR YELRAEWKQP
351 FPGFEALRVH LNRNDYRHDE KAGDAVENFF NNQTQONARIE LRHQPIGRK
401 GSWGVOYLQK KSSALSATSE AVKQPMLLDN KVQHSYFFGV EQANWDNFTL
451 EGGVRVEKQK ASIRYDKALI DRENNYHPL PDLGAHRQTA RSFALSGNWX
501 FTPQHKLST ASHQERLPST QELYAHGKHV ATNTFEVGNK HLNKERSNNI
551 ELALGYEGDR WQYNLALYRN RFGNYIYAQT LNDGRGPKSI EDDSEMKLVR
```

415

601 YNQSGADFYG AEGEIYFKPT PRYRIGVSGD YVRGR LKNLP SLPGRE DAYG
651 NRPLIAQADQ NAPRVPAARL GVHLKASLTD RIDANLDYYR VFAQNKLARY
701 ETRTPGHML NLGANYRRNT RYGEWNWYVK ADNLLNQSVY AHSSFLSDTP
751 QMGRSFTGGV NVKF*

m147/a147 98.1% identity in 734 aa overlap

m147.pep				10	20	30
				PHKTEQSV	DLETVS	VVGKSRPRATSGLLHTS
a147	MRREAKMAQTTLKPIVLSILLINTPLLSQA	HGTEQSV	GLETVS	VVGKSRPRATSGLLHTS		
	10	20	30	40	50	60
m147.pep		40	50	60	70	80
		TASDKIISGDTLRQKAVNLGDALDGVPGI	HASQYGGGASAPVIRGQTGRRIKVLNHHGET			
a147		TASDKIISGDTLRQKAVNLGDALDGVPGI	HASQYGGGASAPVIRGQTGRRIKVLNHHGET			
		70	80	90	100	110
m147.pep						120
m147.pep		100	110	120	130	140
		GDMADFS	PDHAIMVDTALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVS			
a147		GDMADFS	PDHAIMVDSALSQQVEILRGPVTLLYSSGNVAGLVDVADGKIPEKMPENGVS			
		130	140	150	160	170
m147.pep						180
m147.pep		160	170	180	190	200
		ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQ				
a147		ELGLRLSSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSGDYAVPRYRNKRLPDSHADSQ				
		190	200	210	220	230
m147.pep						240
m147.pep		220	230	240	250	260
		TGSIGLSWVGEGKGFIVAYS	DRRDQYGLPAHSHEYDDCHADI	IWQKSLINKRYLQLYPHL		
a147		TGSIGLSWVGEGKGFIVAYS	DRRDQYGLPAHSHEYDDCHADI	IWQKSLINKRYLQLYPHL		
		250	260	270	280	290
m147.pep						300
m147.pep		280	290	300	310	320
		LTEEDIDYDNPGLSCGFHDDDDNAHAH	THSGRPWIDLNRNKRYELRAEWKQFPFGFEALRVH			
a147		LTEEDIDYDNPGLSCGFHDDDDNAHAH	THSGRPWIDLNRNKRYELRAEWKQFPFGFEALRVH			
		310	320	330	340	350
m147.pep						360
m147.pep		340	350	360	370	380
		LNRNDYRHDEKAGDAVENFFN	QTNARIELRHQPIGRLKGSWGVQYLQKSSALSATSE			
a147		LNRNDYRHDEKAGDAVENFFN	QTNARIELRHQPIGRLKGSWGVQYLQKSSALSATSE			
		370	380	390	400	410
m147.pep						420
m147.pep		400	410	420	430	440
		AVKQPMLLDNKVQHYSFFGVEQANWDN	FTLEGGVRVEKQKASIQYDKALIDRENYNHPL			
a147		AVKQPMLLDNKVQHYSFFGVEQANWDN	FTLEGGVRVEKQKASIRYDKALIDRENYNHPL			
		430	440	450	460	470
m147.pep						480
m147.pep		460	470	480	490	500
		PDLGAHRQTARSFALSGN	WYFTTPQHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNK			
a147		PDLGAHRQTARSFALSGN	WYFTTPQHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNK			
		490	500	510	520	530
m147.pep						540
m147.pep		520	530	540	550	560
		HLNKERSNNIELALGYEGDRWQYNLALYRN	RFGNYIYAQT	LNDGRGPKSIEDDSEMKLVR		
a147		HLNKERSNNIELALGYEGDRWQYNLALYRN	RFGNYIYAQT	LNDGRGPKSIEDDSEMKLVR		

416

	550	560	570	580	590	600
m147.pep	580	590	600	610	620	630
	YNQSGADFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFFIAQDDQ					
a147	YNQSGADFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFFIAQADQ					
	610	620	630	640	650	660
m147.pep	640	650	660	670	680	690
	NAPRVPAARLGFHLKASLTDRIANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNT					
a147	NAPRVPAARLGVHLKASLTDRIANLDYYRVFAQNKLARYETRTPGHHMLNLGANYRRNT					
	670	680	690	700	710	720
m147.pep	700	710	720	730		
	RYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
a147	RYGEWNWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
	730	740	750	760		

The following partial DNA sequence was identified in *N. gonorrhoeae* <SEQ ID 589>:

```

g148.seq
1  ATGGCGTTAA  AAACATCAAA  CTGGAACAC  GCAATGctgg  ttcaTCCCGA
51  AgctATgagt  gtcggcgCGC  TTGccgAcaa  AATCCGCAAA  AtcgaAAact
101 gGCCGCAAAA  AGgcaTCTTA  TTCCACGACA  TCACGCCCGT  CCTGCAAAGT
151 GCGGAATACT  TCCGCCTTTT  GGTGCAATTG  CTGGTTTACC  GCTATATGGA
201 TCAGAAAAATC  GACATCGTTG  CCGGCTTGGA  CGCGCGCGGC  TTCATTATCG
251 GCGCGGCACT  CGCCTACCAG  CTCAaCGtcg  gctTCGTCCC  CATCCGCAAA
301 AAAGGCAAGC  TGCCTTTTGA  AACCGTATCG  CAAAGCTAag  cgcTCGAATA
351 CGGGGAAGCT  GCGGTGGAAA  TCCACACCGa  tgcgTCAAA  CCCGGTTCGC
401 GCGTCCTGCT  GGTGATGAT  TTGGTTGCCA  CGGGCGGCAC  AATGCTTGCC
451 GGGCTGGAAC  TGATCCGCAA  ACTCGGCGGG  GAAATTGTCTG  AAgccgccgC
501 CATTTTGGAA  TTTACCGACC  TTCAAGGCGG  CAAGAATATC  CGCGCAAGTG
551 GCGCGCCCTT  ATTTACCCTG  CTTCAAAACG  AAGGCTGCAT  GAAAGGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 590; ORF 148.ng>:

```

g148.pep
1  MALKTSNLEH  AMLVHPEAMS  VGALADKIRK  IENWPQKGIL  FHDITPVLQS
51  AEYFRLLVDL  LVYRYMDQKI  DIVAGLDARG  FIIGAALAYQ  LNVGFVPIRK
101 KGKLPFETVS  QSYALEYGEA  AVEIHTDAVK  PGSRVLLVDD  LVATGGTMLA
151 GLELIRKLGG  EIVEAAAILE  FTDLQGGKNI  RASGAPLFTL  LQNEGCMKG*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 591>:

```

m148.seq
1  ATGGCGTTAA  AAACATCAAA  CTGGAACAC  GCAATGCTGG  TTCATCCCGA
51  AGCTATGAGT  GTCGGCGCGC  TTGCCGACAA  AATCCGCAAA  ATCGAAAAC
101 GGCCGCAAAA  AGGCATCTTA  TTCCACGACA  TCACGCCCGT  CCTTCAAAGC
151 GCGGAATACT  TCCGCCTTTT  GGTGATTTA  TTGGTTTACC  GCTATATGGA
201 TCAGAAAAATC  GACATCGTTG  CCGGTTTGGA  CGCGCGCGGC  TTCATTATCG
251 GCGCGGCACT  CGCCTACCAG  CTCAACGTCG  GTTTCGTCCC  CATCCGCAAA
301 AAAGGCAAGC  TGCCTTTTGA  AACCGTATCG  CAAAGCTACG  CGCTCGAATA
351 CGGGGAAGCT  GCGGTGGAAA  TCCACACCGA  TGCCGTCAAA  CTCGGTTCGC
401 GCGTCTGCT  GGTGATGAT  TTGATTGCCA  CGGGCGGCAC  GATGCTTGCC
451 GGACTGGAAC  TGATCCGCAA  ACTCGGCGGA  GAAATTGTCTG  AAGCCGCCGC
501 CATTTTGGAA  TTTACCGACC  TTCAAGGCGG  CAAGAATATC  CGTGCAAGCG
551 GCGCGCCCTT  ATTTACCCTG  CTTCAAAACG  AAGGCTGTAT  GAAGGCTGA

```

This corresponds to the amino acid sequence <SEQ ID 592; ORF 148>:

```

m148.pep
1  MALKTSNLEH  AMLVHPEAMS  VGALADKIRK  IENWPQKGIL  FHDITPVLQS
51  AEYFRLLVDL  LVYRYMDQKI  DIVAGLDARG  FIIGAALAYQ  LNVGFVPIRK

```

417

101 KGKLPFETVS QSYALEYGEA AVEIHTDAVK LGSRVLLVDD LIATGGTMLA
 151 GLELIRKLGG EIVEAAAILE FTDLQGGKNI RASGAPLFTL LQNEGCMKG*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N. gonorrhoeae*

m148 / g148 99.0% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
g148	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
	10	20	30	40	50	60
	70	80	90	100	110	120
m148.pep	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPPIR	KKGKLPFETV	SQSYALEYGEA
g148	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPPIR	KKGKLPFETV	SQSYALEYGEA
	70	80	90	100	110	120
	130	140	150	160	170	180
m148.pep	AVEIHTDAVK	LGSRVLLVDD	LIATGGTMLA	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI
g148	AVEIHTDAVK	PGSRVLLVDD	LVATGGTMLA	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI
	130	140	150	160	170	180
	190	200				
m148.pep	RASGAPLFTL	LQNEGCMKGX				
g148	RASGAPLFTL	LQNEGCMKGX				
	190	200				

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 593>:

a148.seq

1	ATGGCGTTAA	AAACATCAAA	CTTGGAACAC	GCAATGCTGG	TTCATCCCGA
51	AGCTATGAGT	GTCGGTGCGC	TTGCCGACAA	AATCCGCAAA	ATCGAAAAT
101	GGCCGCAAAA	AGGCATCTTA	TTCCACGACA	TCACGCCCGT	CCTGCAAAGC
151	GCGGAATACT	TCCGACTTTT	GGTTGATTTA	TTGGTTTACC	GCTATATGGA
201	TCAGAAAATC	GACATCGTTG	CCGGTTTGGA	CGCGCGCGGC	TTCATTATCG
251	GCGCGGCACT	CGCCTACCAG	CTCAACGTCG	GTTTCGTCCC	CATCCGCAAA
301	AAAGGCAAGC	TGCCTTTTGA	AACCGTATCG	CAAAGCTACG	CGCTCGAATA
351	CGGGGAAGCT	GCGGTGGAAA	TCCACACCGA	TGCCGTCAAA	CTCGGTTCCG
401	GCGTGCTGCT	GGTCGATGAT	TTGGTTGCCA	CGGCGGGCAC	GATGCTTGCC
451	GGACTGGAGC	TGATCCGCAA	ACTCGGCGGG	GAAATGTGCG	AAGCCGCCGC
501	CATTTTGGA	TTTACCGACC	TTCAAGGCGG	CAAGAATATC	CGTGCAAGCG
551	GCGCGCCCTT	ATTTACCCTG	CTTCAAAACG	AAGGCTGTAT	GAAGGGCTGA

This corresponds to the amino acid sequence <SEQ ID 594; ORF 148.a>:

a148.pep

1	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQS
51	AEYFRLLVDL	LVYRYMDQKI	DIVAGLDARG	FIIGAALAYQ	LNVGFPPIRK
101	KGKLPFETVS	QSYALEYGEA	AVEIHTDAVK	LGSRVLLVDD	LVATGGTMLA
151	GLELIRKLGG	EIVEAAAILE	FTDLQGGKNI	RASGAPLFTL	LQNEGCMKG*

m148/a148 99.5% identity in 199 aa overlap

	10	20	30	40	50	60
m148.pep	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
a148	MALKTSNLEH	AMLVHPEAMS	VGALADKIRK	IENWPQKGIL	FHDITPVLQ	SAEYFRLLVDL
	10	20	30	40	50	60
	70	80	90	100	110	120

```

m148.pep  LVYRYMDQKIDIVAGLDARGFII GAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
a148      LVYRYMDQKIDIVAGLDARGFII GAALAYQLNVGFVPIRKKGKLPFETVSQSYALEYGEA
          70          80          90          100         110         120

          130         140         150         160         170         180
m148.pep  AVEIHTDAVKLGSRVLLVDDLIATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
a148      AVEIHTDAVKLGSRVLLVDDLVATGGTMLAGLELIRKLGGEIVEAAAILEFTDLQGGKNI
          130         140         150         160         170         180

          190         200
m148.pep  RASGAPLFTLLQNEGCMKGX
          |||||||||||||||||||
a148      RASGAPLFTLLQNEGCMKGX
          190         200

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 595>:

```

g149.seq
1  ATGTTGATTG ACAACAATGT CCGCCATTAC AGCTTTTTCG GTGTAGAACA
51  GGCAAATTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATCCGGTAC GACAAAGCAT TGATTGATCG AGAAAACTAC
151 TACAACCAGC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 GTTCGCACTT TCGGGCAACT GGTATTTCAC GCCACACCAC AAACCTCAGCC
251 TGACCGCCTC CCATCAGGAA cgCCTGCCGT CAACGCaagA actGtACgca
301 cacggcAAGC ACGtcgccac CAACACCTTT GAagtccgca acaaACACCT
351 CAACAAAGAG Cgttccaaca atategaACT CGCGCTGGgc tAcaaaggcg
401 accGCTGGCA ATACAATCTG GCAGCCTACC GCAACCGAtT CGGCAACTAC
451 ATTTACGCCC AAACCTTAaa cgacggacgC GGCCCCAAAT CCATCgaaga
501 cgacagcgAA ATGaagcTCG TCGCTACAA CCAATCCGGT GCCGACTTCT
551 ACGgcgcggA aggcgaaatc tACTTcaaaC CGACACCGCG CTACCGCATC
601 GGTGTTTCCG GCGACTatgt acgaggccgT CTGAAAAACC TGCCGTCCCT
651 ACCCGGCAGG gaagatccCT AcggcAAACG TCccttcaTC GCACAAGCCG
701 ACCAAAACGC CCCC CGCATT ceggetGCGC GCCTCGGCTT CCACCTGAAA
751 ACCTCGCTAA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCAAAAC AAACCTCGCC GCTACGAAAC GCGTACGCCC GGACACCATA
851 TGCTCAACCT CGGTGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCtc aACcaatCcg tTTACGCCCa
951 cAGCAGCTTC CTCTCTGATA CGCCGCAAAc gGGCCGCAGC TTtgccgGCg
1001 gcgtaAACGT GaAGTTttaA

```

This corresponds to the amino acid sequence <SEQ ID 596; ORF 149.ng>:

```

g149.pep
1  MLIDNNVRHY SFFGVEQANW DNFTLEGGVR VEKQKASIRY DKALIDRENY
51  YNQPLPDLGA HRQTARSPAL SGNWYFTPHH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YKGDWRQYNL AAYRNRFGNY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVSGDYVRGR LKNLPSLPGR EDPYGKRPFi AQADQNAPRI PAARLGFHLK
251 TSLTDRIDAN LDYYRVFAQN KLARYETRTP GHHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FAGGVNVKF*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 597>:

```

m149.seq
1  ATGCTGCTTG ACAACAAAGT GCAACATTAC AGCTTTTTCG GTGTAGAACA
51  GGCAAATGGG GACAACTTCA CGCTTGAAGG AGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATTCAGTAC GACAAAGCAT TGATTGATCG GGAAAACTAC
151 TACAACCACC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCCGCTC
201 ATTCGCACTT TCGGGCAACT GGTATTTCAC GCCACAACAC AAACCTCAGCC
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA
301 CACGGCAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAAACACCT
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGGCG
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGTAACCTAC

```



```

451 ATTTACGCCC AAACCTTAAA CGACGGACGC GGCCCCAAAT CCATCGAAGA
501 CGACAGCGAA ATGAAGCTCG TGCGCTACAA CCAATCCGGC GCCGACTTCT
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
601 GCGGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAAACC TGCCTTCCCT
651 ACCCGGCAGA GAAGATGCCT ACGGCAACCG TCCTTTCATC GCACAGGACG
701 ACCAAATATG CCCCCGTGTT CCGGCTGCGC GCCTCGGCTT CCACCTGAAA
751 GCCTCGCTGA CCGACCGTAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCAAAAC AAACCTCGCCC GCTACGAAAC GCGCACGCCC GGACACCATA
851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
951 CAGCAGCTTT CTCTCTGATA CGCCGCAAAT GGGCCGCGC TTTACCGCGC
1001 GCGTGAACGT GAAGTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 598; ORF 149>:

m149.pep

```

1 MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY
51 YNHPLPDLGA HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFNGY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVSVDYVRGR LKNLPSLPGR EDAYGNRPFI AQDDQNAPRV PAARLGFHLK
251 ASLTDRIAN LDYYRVFAQN KLARYETRT PGHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 149 shows 95.9% identity over a 339 aa overlap with a predicted ORF (ORF 149.ng) from *N. gonorrhoeae*:

m149/g149

	10	20	30	40	50	60
m149.pep	MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIQY DKALIDRENY YNHPLPDLGA					
	: : :					
g149	MLIDNNVRHYSFFGVEQANW DNFTLEGGVR VEKQKASIRYDKALIDRENY NQPLPDLGA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m149.pep	HRQTARFALSGNWYFTPQH KLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
	:					
g149	HRQTARFALSGNWYTFPHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
	70	80	90	100	110	120
	130	140	150	160	170	180
m149.pep	RSNNIELALGYEGDRWQYNLALYRNRFNGYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
	:					
g149	RSNNIELALGYKGDWQYNLAAYRNRFNGYIYAQTLNDGRGPKSIEDDSEMKLVRYNQSG					
	130	140	150	160	170	180
	190	200	210	220	230	240
m149.pep	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGR LKNLPSLPGREDAYGNRPFI AQDDQNAPRV					
	: :					
g149	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGR LKNLPSLPGREDPYGKRPFIAQADQNAPRI					
	190	200	210	220	230	240
	250	260	270	280	290	300
m149.pep	PAARLGFHLKASLTDRIANLDYYRVFAQNKLARYETRTPGHMLNLGAN YRRNTRYGEW					
	:					
g149	PAARLGFHLKSLTDRIANLDYYRVFAQNKLARYETRTPGHMLNLGAN YRRNTRYGEW					
	250	260	270	280	290	300
	310	320	330	340		
m149.pep	NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX					
	:					

420

g149 NWYVKADNLLNQSVYAHSSFLSDTPQMGRSFAGGVNVKFX
310 320 330 340

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 599>:

a149.seq
1 ATGCTGCTTG ACAATAAAGT GCAACATTAC AGCTTTTTCG GTGTAGAACA
51 GGCAAACCTGG GACAACTTCA CGCTTGAAGG CGGCGTACGC GTGGAAAAAC
101 AAAAAGCCTC CATCCGCTAC GACAAAGCAT TGATTGATCG GGAAACTAC
151 TACAACCATC CCCTGCCCGA CCTCGGCGCG CACCGCCAAA CCGCCGCTC
201 ATTCGCACTT TCGGGCAACT GGTATTTCAC GCCACAACAC AAACCTAGCC
251 TGACCGCCTC CCATCAGGAA CGCCTGCCGT CAACGCAAGA GCTGTACGCA
301 CACGGCAAAC ACGTCGCCAC CAACACCTTT GAAGTCGGCA ACAAACACCT
351 CAACAAAGAG CGTTCCAACA ATATCGAACT CGCGCTGGGC TACGAAGCGG
401 ACCGCTGGCA ATACAATCTG GCACTCTACC GCAACCGCTT CGGCAACTAC
451 ATTTACGCC AAACCTTAAA CGACGGACGC GGCCCAAAAT CCATCGAAGA
501 CGACAGCGAA ATGAAGCTCG TCGCTACAA CCAATCCGGT GCGGACTTCT
551 ACGGCGCGGA AGGCGAAATC TACTTCAAAC CGACACCGCG CTACCGCATC
601 GCGGTTTCCG GCGACTATGT ACGAGGCCGT CTGAAAACC TGCCTTCCCT
651 ACCCGGCAGG GAAGACGCCT ACGGCAACCG CCCACTCATT GCCCAAGCCG
701 ACCAAAACGC CCCTCGGCTT CCGGCTGCGC GCCTCGGCGT CCACCTGAAA
751 GCCTCGCTGA CCGACCGCAT CGATGCCAAT TTGGACTACT ACCGCGTGTT
801 CGCCCAAAAC AAACCTGCCC GCTACGAAAC GCGCACGCCC GGACACCATA
851 TGCTCAACCT CGGCGCAAAC TACCGCCGCA ATACGCGCTA TGGCGAGTGG
901 AATTGGTACG TCAAAGCCGA CAACCTGCTC AACCAATCCG TTTACGCCCA
951 CAGCAGCTTC CTCTCTGATA CGCCGCAAAT GGGCCGCAGC TTTACCGGCG
1001 GCGTGAACGT GAAGTTTAA

This corresponds to the amino acid sequence <SEQ ID 600; ORF 149.a>:

a149.pep
1 MLLDNKVQHY SFFGVEQANW DNFTLEGGVR VEKQKASIRY DKALIDRENY
51 YNHPLPDLGA HRQTARFAL SGNWYFTPQH KLSLTASHQE RLPSTQELYA
101 HGKHVATNTF EVGNKHLNKE RSNNIELALG YEGDRWQYNL ALYRNRFNGY
151 IYAQTLNDGR GPKSIEDDSE MKLVRYNQSG ADFYGAEGEI YFKPTPRYRI
201 GVS GDYVRGR LKNLPSLPGR EDAYGNRPLI AQADQNAPRV PAARLGVLK
251 ASLTDRIDAN LDYYRVFAQN KLARYETRTP GHMLNLGAN YRRNTRYGEW
301 NWYVKADNLL NQSVYAHSSF LSDTPQMGRS FTGGVNVKF*

m149/a149 98.8% identity in 339 aa overlap

	10	20	30	40	50	60
m149.pep	MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPLPDLGA					
a149	MLLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPLPDLGA					
	10	20	30	40	50	60
	70	80	90	100	110	120
m149.pep	HRQTARFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
a149	HRQTARFALSGNWYFTPQHKLSLTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKE					
	70	80	90	100	110	120
	130	140	150	160	170	180
m149.pep	RSNNIELALGYEGDRWQYNLALYRNRFNGYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSG					
a149	RSNNIELALGYEGDRWQYNLALYRNRFNGYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSG					
	130	140	150	160	170	180
	190	200	210	220	230	240
m149.pep	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGLKNLPSLPGREDAYGNRPFIAQDDQNAPRV					
a149	ADFYGAEGEIYFKPTPRYRIGVSGDYVRGLKNLPSLPGREDAYGNRPLIAQADQNAPRV					
	190	200	210	220	230	240
	250	260	270	280	290	300
m149.pep	PAARLGFLKASLTDRIDANLDYYRVFAQNKLARYETRTPGHMLNLGAN YRRNTRYGEW					

g149-1.seq

1	ATGGCACAAA	TCACACTCAA	ACCCATTGTT	TTATCAATTC	TTTTAATCAA
51	CACACCCCTC	CTCGCCCAAG	CGCATGAAAC	TGAGCAATCG	GTGGGCTTGG
101	AAACGGTCA	CGTCGTCGGC	AAAAGCCGTC	CGCGCGCATG	TTCCGGGGCTG
151	CTGCACACTT	CAGCCGGCTG	CGACAAAATC	ATCTCCGGCG	ATACTTTGCG
201	CCAAAAAGCC	GTCAACTTGG	GGGACGCTAT	GGACGGCGTA	CCGGGCATCC
251	ACGCTTCCGA	ATACGGCGCG	GGCGCATCCG	C7CCCGTTAT	TCGGCGGTCAA
301	ACGGGACAG	GGATTAAAGT	ATTGAACCAT	CACGGCGAAA	CGGGCGATAT
351	GGCGGACTTT	TCTCCCGATC	ACGCCATTAT	GGTAGATACC	GCCTTGTCGG
401	ACAAGGTTGA	ATCTCTGCGC	GGGCGGGTTA	CGCTCTTGTA	CAGCTTCGGCG
451	AATGTGGCGG	GAGTGTCGCG	TGTTGGCGAT	GGA AAAATTC	CCGAAAAAAT
501	GCCTGAAAAC	GGCGTATCGG	GCGaagccgG	ATTGCGTTTG	AGACGGCGCA
551	ATTTGAAAAA	ACTCGAATCC	GCAGGGATCA	ATATCGGACT	GGGCAAAAAAC
601	TTCTGTGCTG	ATACCGAAGC	CTTGTACCCG	AAATCGGGCG	ATTACGCCGT
651	ACCGCGTTAC	CGCAATCTGA	AAACGCTGCC	GCACAGCCAT	CGCGATTTCG
701	AAACGGGCGC	CATCGGGGCT	TCTTGGGTGG	GGCAAAAAGG	CTTTATCGGC
751	CGACCATACA	GCGACCGCTG	CGACCGCTAC	GGCCTGCGCT	CCCAACAGCA
801	CGAATACGAT	GAGTCGCCAG	CGCAGCATAT	TGCGCAAAAG	AGTTTGATCA
851	ACAAACGCTA	TTTGACGCTT	TATCCGCACT	TGTTGACCGA	AGAAGACATC
901	GATTACGCA	ATCCGGGGCT	GAGCTCGCGG	T7CCACGACG	CGCAGCGGTGC
951	ACACGCACAC	ACCCACAACG	GCAAAACGTF	GATAGACTCG	CGCAACAACAT
1001	GCTACGAACT	CCGCGCCGAA	TGGAAGCAGC	CATTCCCCGG	TTTTGAAGCC
1051	CTGCGCGTAT	ATCTGAACCT	CAATGACTAC	CACCACGACG	AAAAGACAGG
1101	CGATGCAGTA	GAAAATCTCT	TCAACACAAC	AACACACAC	GGCCGATATC
1151	AGTTGCGCCA	CCAACCCATA	GGCCGTCTGA	AAGGCAGCTG	GCGCCGTGCA
1201	TATTTGGGAC	AAAAATCGAC	CGCGCTTCCC	GCCATTCCCG	AAACCGTGCA
1251	ACAACCGATG	TTGATTGACA	ACAATTGCTC	CCATTACAGC	TTTTTCGGTG
1301	TAGAACACGG	ACATTGGGAC	AACTTACACG	TGAAGCGCGG	CGTACGCGTG
1351	GAAAAACAAA	ARGCTTCCAT	CGCGTACGAC	AAAGCATTTA	TTGATCGAGA
1401	AACTACTACT	AACCAAGCCC	TGCCGACCTC	CGGCGCGCAT	CGCCAAACCG
1451	CCCGCTCGTT	CGCATTCTCG	GGCAACTGAT	ATTTCACGCC	ACACCCACAA
1501	CTCAGCGTGA	CGCGCTCCCA	TCAGGAACGC	CTGCCGTCAA	CGCAAGAACT
1551	GTACGCACTC	GGCAAGCACG	TGCGCACCAA	CACCTTTGAA	TGCGGCAACG
1601	AACACTCTCA	CAAGAGAGCG	TCCAACAATA	TGCAACTCGC	GCTGGGCTAC
1651	GAAGGGGACC	GCTGGCAATA	CAATCTGGCA	GCCTACCGCA	ACCGATTTCG
1701	CRACTACATT	TACGCCCCAA	CCTTAAACGA	CGGACGCGGC	CCCAAAATCCA
1751	TGCAAGACGA	CAGGCAAAAT	AAGCTCTGCG	GCTACAACCA	ATCCGGTGCT
1801	GACTTCTACG	GCGCGGAAGG	CGAAATCTAC	TTCAAACCGA	CACCGCGCTA
1851	CGCATCTCGT	GTTTCCGGCG	ACTATGTACG	AGGCCGCTCG	AAAAACCTGC
1901	CGTCCCTACT	GGGACGGGAA	GATCCCTACG	GCAAAACGTC	CTTCATCGCA
1951	CAAGCCGACC	AAAACGCCCC	CGCATTTCCG	GTCGCGCGCC	TCGGGTTCCA
2001	CCTGAAAAAC	TCGCTAACCG	ACCGTATCGA	TGCCAATTTG	GACTACTACC
2051	GCGTGTTCCG	CCAAAACAAA	CTCGCCCGCT	ACGAAACGCG	TACGCGCGGA
2101	ACCATATATG	TCAACCTCGG	TGCAAACTAC	CGCCGCAATA	CGCGCTATGG
2151	CGAGTGGAA	TGGTAGCTCA	AAGCGACAAA	CCTGCTCAAC	CAATCCGTTT
2201	ACGCCACACG	CAGCTTCTCT	TCTGATACCG	CGCAAAATGG	CCGACGTTTT
2251	ACCGCGGGCG	TAAACGTGAA	GTTTTAA		

g149-1.pap

1	MAQITLKPIV	LSILLINTPL	LAQAHETEQS	VGLETSVVVG	KSRPRATSGL
51	LHTSTASDKI	ISGDTLRQKA	VNLGDALDVT	PGIHASQOYGG	GASAPVIRGQ
101	TGRRIKVLNH	HGETGDMADF	SPDHGAIMDV	ALSQOQVEILR	GPVTLLYSYG
151	NVAGLVDDAV	GKIPKMPFN	GVSGEAGLRL	SSGNLEKLTJ	AGINGLGLKN
201	FVLHTEGLRY	KSGDYAVPRY	RNLKRLPDSH	ADSQTSIGSL	SWVGEKGFIG
251	AAYSDDRRDY	GLPAHSHEYD	DCHADTIWQK	SLINKRYLQL	YPHLITTEEDI
301	DYDNPGLSCG	FHDGGDAHAH	THKGPIWIDL	RNKRYELREA	WKOPFFGEA

```

351  LRVLNLRNDY HHDEKAGDAV ENFFNNKTHN ARIELRHQPI GRLKGSWGVQ
401  YLGQKSSALS AIPETVQOPM LIDNNVRHYS FFGVEQANWD NFTLEGGVRV
451  EKQKASIRYD KALIDRENY NQPLPDLGAH RQTARSFALS GNWYFTPHHK
501  LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY
551  EGDWRQYNLA AYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
601  DFYGAEGEYI FKPTPRYRIG VSGDYVRGRL KNLPSLPGRE DPGYKRPPIA
651  QADQNAPRIP AARLGPHLKT SLTDRIDANL DYYRVFAQNK LARYETRTPG
701  HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
751  TGGVNVKF*

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 603>:

```

m149-1.seq
1  ATGGCACAAA CTACACTCAA ACCCATGTGT TTATCAATTC TTTTAATCAA
51  CACACCCCTC CTGCCCCAAG CGCATGAAAC TGAGCAATCG GTGGATTGG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGCCAC GTCGGGGCTG
151 TTGCACACTT CGACCGCCTC CGACAAAATC ATCTCCGGCG ATACCTTGCG
201 CCAAAAGGCC GTCAACTTGG GCGACGCTTT AGACGGCGTA CCGGGCATCC
251 ACGCTTCGCA ATACGGCGGC GCGCGCTCTG CTCCCGTCAT TCGCGGTCAA
301 ACAGGCAGGC GGATTAAAGT GTTGAACCAT CACGGCGAAA CAGGCGATAT
351 GCGGGATTTT TCGCCCGATC ACGCCATTAT GGTAGATACC GCCTTGTCGC
401 AACAGTTCGA AATCTGCGC GGGCCGGTTA CGCTCTTGTA CAGCTCGGGC
451 AATGTGGCGG GGCTGGTCGA TGTGCGCAT GGCAAAATCC CCGAAAAAAT
501 GCCTGAAAAC GCGGTATCGG GCGAACTCGG ATTGCGTTTG AGCAGCGGCA
551 ATCTGGAAAA ACTCAGCTCC GCGCGCATCA ATATCGGTTT GGGCAAAAAC
601 TTGTGATTGC ACACGGAAGG GCTGTACCGC AAATCGGGGG ATTACGCCGT
651 ACCGCGTTAC CGCAATCTGA AACGCTGCGC CGACAGCCAC GCCGATTGCG
701 AAACGGGCGA CATCGGGCTG TCTTGGGTTG GCGAAAAAGG TTTTATCGGC
751 GTAGCGTACA GCGACCGTCG CGACCAATAT GGTCTGCGTG CCCACAGCCA
801 CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGCTTGATTA
851 ACAACGCTA TTTACAGCTT TATCCGCACC TGTTGACCGA AGAAGACATC
901 GATTACGACA ATCCGGGCTT GAGCTGCGGC TTCCACGACG ACGATAATGC
951 ACACGCACAC ACCCACAGCG GCAGACCGTG GATAGACCTG CGCAACAAAC
1001 GCTACGAACT CCGTGCCGAA TGAAGCAAC CGTTCGCCGG TTTTGAAGCC
1051 CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1101 CGATGCAGTC GAAAACCTTT TTAACAACCA AACGCAAAAC GCCCGCATCG
1151 AGTTGCGCCA CCAACCCATA GGTCTGCTGA AAGGCAGCTG GGGCGTGCAA
1201 TATTTACAAC AAAAATCCAG TGCTTTATCT GCCATATCCG AAGCGGTTAA
1251 ACAACCGATG CTGCTTGACA ACAAAAGTGA ACATTACAGC TTTTTCGGTG
1301 TAGAACAGGC AAACCTGGAC AACTTCACGC TTGAAGGAGG CGTACGCGTG
1351 GAAAAACAAA AAGCCTCCAT TCAGTACGAC AAAGCATTGA TTGATCGGGA
1401 AAACCTACTAC AACCACCCCT TGCCCGACCT CGGCGCGCAC GCACAAACCG
1451 CCCGCTCAAT CGCACTTTTCG GGCAACTGGT ATTTACGCCC ACAACACAAA
1501 CTCAGCCTGA CCGCCTCCCA TCAGGAACGC CTGCCGTCAG CGCAAGAGCT
1551 GTACGCACAC GGCACACACG TCGCCACCAA CACCTTTGAA TCGCGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
1651 GAAGGCGACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1701 TAACCTACATT TACGCCCCAA CCTTAAACGA CGGACGCGGC CCAAAATCCA
1751 TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGCGCC
1801 GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1851 CCGCATCGGC GTTTCGGCGC ACTATGTACG AGGCGGCTCG AAAAACCTGC
1901 CTTCCCTACC CGGCAGAGAA GATGCCTACG GCAACCGTCC TTTCATCGCA
1951 CAGGACGACC AAAATGCCCC CCGTGTTCGG GCTGCGCGCC TCGGCTTCCA
2001 CCGTAAAGCG TCGCTGACCG ACCGTATCGA TGCCAATTG GACTACTACC
2051 GCGTGTTCGC CAAAACAAA CTCGCCCGCT ACGAAACGCG CACGCCCGGA
2101 CACCATATGC TCAACCTCGG CGCAAACTAC CGCCGCAATA CGCGCTATGG
2151 CGAGTGGAAT TGGTACGTCA AAGCCGACAA CCGTCTCAAC CAATCCGTTT
2201 ACGCCACAG CAGCTTTCTC TCTGATACGC CGCAAATGGG CCGCAGCTTT
2251 ACCGGCGGCG TGAACGTGAA GTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 604; ORF 149-1>:

```

m149-1.pep
1  MAQTLKPIV LSILLINTPL LAQAHETEQS VDLETVSVMG KSRPRATSGL
51  LHTSTASDKI ISGDTLRQKA VNLGDALDGV PGIHASQYGG GASAPVIRGQ
101 TGRRIKVLNH HGETGDMADF SPDHAIMVDT ALSQQVEILR GPVTLLYSSG
151 NVAGLVDVAD GKIPKMPEN GVSGLGLRL SSGNLEKITS GGINIGLGKN
201 FVLHTEGLYR KSGDYAVPRY RNLKRLPDSh ADSQTGSIGL SWVGEKGFIF
251 VAYSDDRDQY GLPAHSHEYD DCHADIIWQK SLINKRYLQL YPHLLTEEDI
301 DYDNFGLSCG FHDDDNAHAH THSGRPWIDL RNKRYELRAE WKQFPFGFEA
351 LRVLNLRNDY RHDEKAGDAV ENFFNNQTN ARIELRHQPI GRLKGSWGVQ
401 YLGQKSSALS AISEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRV
451 EKQKASIQYD KALIDRENY NHPPLDLGAH RQTARSFALS GNWYFTPHHK
501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKER SNNIELALGY

```

423

551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
 601 DFYGAEGEY FKPTPRYRIG VSGDYVRGRL KNLPSPGRE DAYGNRPFFIA
 651 QDDQNAAPRV AARLGFLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG
 701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
 751 TGGVNVKF*

m149-1/g149-1 96.2% identity in 758 aa overlap

	10	20	30	40	50	60
m149-1.pep	MAQTTLKPIVLSILLINTPLLAQAHETEQSVQLETVSVVGKSRPRATSGLLHTSTASDKI					
g149-1	MAQTTLKPIVLSILLINTPLLAQAHETEQSVQLETVSVVGKSRPRATSGLLHTSTASDKI					
	10	20	30	40	50	60
	70	80	90	100	110	120
m149-1.pep	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
g149-1	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
	70	80	90	100	110	120
	130	140	150	160	170	180
m149-1.pep	SPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVDVADGKIPEKMPENGVSSELGLRL					
g149-1	SPDHAIMVDTALSQQVEILRGPVTLTYSSGNVAGLVDVADGKIPEKMPENGVSSELGLRL					
	130	140	150	160	170	180
	190	200	210	220	230	240
m149-1.pep	SSGNLEKLTSGGINIGLGNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL					
g149-1	SSGNLEKLTSGGINIGLGNFVLHTEGLYRKSGDYAVPRYRNLKRLPDSHADSQTGSIGL					
	190	200	210	220	230	240
	250	260	270	280	290	300
m149-1.pep	SWVGEKGFIVAYSDRRDQYGLPAHSHEYDDCHADIWQKSLINKRYLQLYPHLLTEEDI					
g149-1	SWVGEKGFIVAYSDRRDQYGLPAHSHEYDDCHADIWQKSLINKRYLQLYPHLLTEEDI					
	250	260	270	280	290	300
	310	320	330	340	350	360
m149-1.pep	DYDNPGLSGCGFHDHDDNAHAHTHSGRPWIDLNRKRYELRAEWKQFPFGFEALRVHLNRNDY					
g149-1	DYDNPGLSGCGFHDHDDNAHAHTHSGRPWIDLNRKRYELRAEWKQFPFGFEALRVHLNRNDY					
	310	320	330	340	350	360
	370	380	390	400	410	420
m149-1.pep	RHDEKAGDAVENFFNNQTONARIELRHQPIGRLKGSWGVQYLQKSSALSASEAVKQPM					
g149-1	HHDEKAGDAVENFFNNKTHNARIELRHQPIGRLKGSWGVQYLQKSSALSASEAVKQPM					
	370	380	390	400	410	420
	430	440	450	460	470	480
m149-1.pep	LLDNKVQHYSEFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYNHPPLDLGAH					
g149-1	LIDNNVRHYSEFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPPLDLGAH					
	430	440	450	460	470	480
	490	500	510	520	530	540
m149-1.pep	RQTARSAFALSGNWFYTPQHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKR					
g149-1	RQTARSAFALSGNWFYTPQHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKR					
	490	500	510	520	530	540
	550	560	570	580	590	600
m149-1.pep	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMCLVRYNQSGA					
g149-1	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMCLVRYNQSGA					
	550	560	570	580	590	600
	610	620	630	640	650	660
m149-1.pep	DFYGAEGEYFKPTPRYRIGVSGDYVRGRLKNLPSPGREDAYGNRPFFIAQDDQNAAPRV					
g149-1	DFYGAEGEYFKPTPRYRIGVSGDYVRGRLKNLPSPGREDFYGRPFIAQADQNAAPRI					
	610	620	630	640	650	660

424

```

      670      680      690      700      710      720
m149-1.pep  AARLGFLKASLTDRIDANLDYRVFAQNKLYETRTPGHHMLNLGANYRRNTRYGEWN
             |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g149-1      AARLGFLKTSLTDRIDANLDYRVFAQNKLYETRTPGHHMLNLGANYRRNTRYGEWN
             670      680      690      700      710      720

      730      740      750      759
m149-1.pep  WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
             |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
g149-1      WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
             730      740      750

```

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 605>:

a149-1.seq

```

1  ATGGGCACAAA CTACACTCAA ACCCATGTGT TTATCAATTC TTTTAATCAA
51  CACACCCCTC CTCTCCCAAG CGCATGGAAC TGAGCAATCA GTGGGCTTGG
101 AAACGGTCAG CGTCGTCGGC AAAAGCCGTC CGCGCGCCAC TTCGGGGCTG
151 CTGCACACTT CTACCGCCTC CGACAAAATC ATCAGCGCGC ACACCTTGCG
201 ACAAAAAGCC GTCAACTTGG GTGATGCTTT AGACGGCGTA CCGGGCATTC
251 ATGCCCTCGA ATACGGCGGC GCGGCATCCG CTCCCGTTAT TCGCGGTCAA
301 ACAGGCAGAC GGATTAAAGT GTTGAACCAT CACGGCGAAA CGGGCGACAT
351 GCGGGACTTC TCTCCAGACC ATGCAATCAT GGTGACAGC GCCTTGTCGC
401 AACAGGTCGA AATCCTGCGC GGTCCGGTTA CGCTCTTGTA CAGCTCGGGC
451 AATGTGGCGG GGCTGGTCGA TGTGCCGAT GGCAAAATCC CCGAAAAAAT
501 GCCTGAAAAC GCGGTATCGG GCGAACTCGG ATTGCGTTTG AGCAGCGGCA
551 ATCTGGAAAA ACTCAGTCC GCGGCATCA ATATCGGTTT GGGCAAAAAC
601 TTTGTATTGC ACACGGAAGG GCTGTACCGC AAATCGGGG ATTACGCCGT
651 ACCGCGTTAC CGCAATCTGA AACGCTGCC CGACAGCCAC GCCGATTCCG
701 AAACGGGCAG CATCGGGCTG TCTTGGGTTG GCGAAAAAGG CTTTATCGGC
751 CGAGCATACA GCGACCGTCG CGACCAATAT GGTCTGCCTG CCCACAGCCA
801 CGAATACGAT GATTGCCACG CCGACATCAT CTGGCAAAAG AGTTTGATTA
851 ACAACGCTA TTTGAGCTT TATCCGCAAC TGTGACCGA AGAAGACATC
901 GATTACGACA ATCCGGGCTT GAGCTGCGGC TTTCACGACG ACGATGATGC
951 ACACGCCCAT GCCCACAACG GCAAACCTTG GATAGACCTG CGCAACAAAC
1001 GCTACGAACT CGCGCGCGAA TGGAAAGCAAC CGTTCGCCGG TTTTGAAGCC
1051 CTGCGCGTAC ACCTGAACCG CAACGACTAC CGCCACGACG AAAAAGCAGG
1101 CGATGCAGTA GAAAACCTTT TTAACAACCA AACGCAAAAC GCCCGTATCG
1151 AGTTGCGCCA CCAACCCATA GGCGGTCTGA AAGGCAGCTG GGGCGTGCAA
1201 TATTTGGGAC AAAAATCCAG TGCTTTATCT GCCACATCCG AAGCGGTCAA
1251 ACAACCGATG CTGCTTGACA ATAAAGTGCA ACATTACAGC TTTTTCGGTG
1301 TAGAACAGGC AAACCTGGGAC AACTTCACGC TTGAAGGCGG CGTACGCGTG
1351 GAAAAACAAA AAGCCTCCAT CCGCTACGAC AAAGCATTGA TTGATCGGGA
1401 AAACACTAC AACCATCCCC TGCCCGACCT CGGCGCGCAC CGCCAAACCG
1451 CCGGCTCATT GCACTTTTCG GGCAACTGGT ATTTACGCC ACAACACAAA
1501 CTCAGCCTGA CCGCTCCCA TCAGGAACGC CTGCCGTCAA CGCAAGAGCT
1551 GTACGCACAG GGCAACACG TCGCCACCAA CACCTTTGAA GTCGGCAACA
1601 AACACCTCAA CAAAGAGCGT TCCAACAATA TCGAACTCGC GCTGGGCTAC
1651 GAAGGCGACC GCTGGCAATA CAATCTGGCA CTCTACCGCA ACCGCTTCGG
1701 CAACTACATT TACGCCCAA CCTTAAACGA CGGACGCGGC CCCAAATCCA
1751 TCGAAGACGA CAGCGAAATG AAGCTCGTGC GCTACAACCA ATCCGGTGCG
1801 GACTTCTACG GCGCGGAAGG CGAAATCTAC TTCAAACCGA CACCGCGCTA
1851 CCGCATCGGC GTTTCGGGCG ACTATGTACG AGGCCGCTG AAAAACCTGC
1901 CTTCCCTACC CGGCAGGGA GACGCTACG GCAACGCCCC ACTCATTGCC
1951 CAAGCCGACC AAAACGCCCC TCGCGTTCCG GCTGCGCGCC TCGGCGTCCA
2001 CCTGAAAGCC TCGTGACCG ACCGCATCGA TGCCAATTTG GACTACTACC
2051 GCGTGTTCG CCAAACAAA CTCGCCCGCT ACGAAACGCG CACGCCCGGA
2101 CACCATATGC TCAACCTCGG CGCAAACTAC CGCGCAATA CGCGCTATGG
2151 CGAGTGGAA TGGTACGTCA AAGCCGACAA CCTGCTCAAC CAATCCGTTT
2201 ACGCCACAG CAGCTTCCTC TCTGATACGC CGCAATGGG CCGCAGCTTT
2251 ACCGCGCGG TGAACGTGAA GTTTTAA

```

This corresponds to the amino acid sequence <SEQ ID 606; ORF 149-1.a>:

a149-1.pep

```

1  MAQTLKPIV LSILLINTPL LSQAHGTEQS VGLETVSVVG KSRPRATSGL
51  LHTSTASDKI ISGDTLRQKA VNLGDALDGV PGIHASQYGG GASAPVIRGQ
101 TGRRIKVLNH HGETGDMADF SPDHAIMVDS ALSQQVEILR GPVTLTYSSG
151 NVAGLVDVAD GKIPEKMPEN GVSSELGLRL SSGNLEKLTG GGINIGLGKN
201 FVLHTEGLYR KSGDYAVPRY RNLKRLPDSH ADSQTGSIGL SWVGEKGFIG
251 AAYSDDRDQY GLPAHSHEYD DCHADIWQK SLINKRYLQL YPHLLTEEDI
301 DYDNFGLSCG FHDDDDAHAH AHNGKPWIDL RNKRYELRAE WKQFPFGFEA

```

351 LRVHLNRNDY RHDEKAGDAV ENFFNNQTON ARIELRHQPI GRLKGSWGVO
 401 YLGQKSSALS ATSEAVKQPM LLDNKVQHYS FFGVEQANWD NFTLEGGVRV
 451 EKQKASIRYD KALIDRENY NHPPLDLGAH RQTARSFALS GNWYFTPOHK
 501 LSLTASHQER LPSTQELYAH GKHVATNTFE VGNKHLNKR SNNIELALGY
 551 EGDWRQYNLA LYRNRFGNYI YAQTLNDGRG PKSIEDDSEM KLVRYNQSGA
 601 DFYGAEGEYI FKPTPRYRIG VSGDYVRGRL KNPLSLPGRE DAYGNRPLIA
 651 QADQNAAPRV AARLGVHLKA SLTDRIDANL DYYRVFAQNK LARYETRTPG
 701 HHMLNLGANY RRNTRYGEWN WYVKADNLLN QSVYAHSSFL SDTPQMGRSF
 751 TGGVNVKF*

a149-1/m149-1 98.0% identity in 758 aa overlap

	10	20	30	40	50	60
a149-1.pep	MAQTTLKPIVLSILLINTPLLSQAHGTEQSVGLETVSVVGKSRPRATSGLLHTSTASDKI					
m149-1	MAQTTLKPIVLSILLINTPLLAQAHETEQSVLDLETVSVVGKSRPRATSGLLHTSTASDKI					
	10	20	30	40	50	60
	70	80	90	100	110	120
a149-1.pep	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
m149-1	ISGDTLRQKAVNLGDALDGVPGIHASQYGGGASAPVIRGQTGRRIKVLNHHGETGDMADF					
	70	80	90	100	110	120
	130	140	150	160	170	180
a149-1.pep	SPDHAIMVDSALSQQVEILRGPVTTLLYSSGNVAGLVDVADGKIPEKMPENGVSSELGLRL					
m149-1	SPDHAIMVDTALSQQVEILRGPVTTLLYSSGNVAGLVDVADGKIPEKMPENGVSSELGLRL					
	130	140	150	160	170	180
	190	200	210	220	230	240
a149-1.pep	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSQDYAVPRYRNLKRLPDSHADSTGTSIGL					
m149-1	SSGNLEKLTSGGINIGLGKNFVLHTEGLYRKSQDYAVPRYRNLKRLPDSHADSTGTSIGL					
	190	200	210	220	230	240
	250	260	270	280	290	300
a149-1.pep	SWVGEKGFIGAAYSDDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI					
m149-1	SWVGEKGFIVAYSDDRRDQYGLPAHSHEYDDCHADIIWQKSLINKRYLQLYPHLLTEEDI					
	250	260	270	280	290	300
	310	320	330	340	350	360
a149-1.pep	DYDNPGLSGCFHDDDDAHAAHNGKPWIDLNRKRYELRAEWKQFPFGFEALRVHLNRNDY					
m149-1	DYDNPGLSGCFHDDDDAHAAHNGKPWIDLNRKRYELRAEWKQFPFGFEALRVHLNRNDY					
	310	320	330	340	350	360
	370	380	390	400	410	420
a149-1.pep	RHDEKAGDAVENFFNNQTONARIELRHQPIGRLKGSWGQYLGQKSSALSATSEAVKQPM					
m149-1	RHDEKAGDAVENFFNNQTONARIELRHQPIGRLKGSWGQYLGQKSSALSATSEAVKQPM					
	370	380	390	400	410	420
	430	440	450	460	470	480
a149-1.pep	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIRYDKALIDRENYNHPPLDLGAH					
m149-1	LLDNKVQHYSFFGVEQANWDNFTLEGGVRVEKQKASIQYDKALIDRENYNHPPLDLGAH					
	430	440	450	460	470	480
	490	500	510	520	530	540
a149-1.pep	RQTARSFALSGNWFYFTPOHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKR					
m149-1	RQTARSFALSGNWFYFTPOHKLSTASHQERLPSTQELYAHGKHVATNTFEVGNKHLNKR					
	490	500	510	520	530	540
	550	560	570	580	590	600
a149-1.pep	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSGA					
m149-1	SNNIELALGYEGDRWQYNLALYRNRFGNYIYAQTLNDGRGPKSIEDDSEMMLVRYNQSGA					
	550	560	570	580	590	600
	610	620	630	640	650	660

426

```

a149-1.pep  DFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPLIAQADQNAAPRVP
|||||
m149-1      DFYGAEGEIIYFKPTPRYRIGVSGDYVRGRLKNLPSLPGREDAYGNRPFIAQDDQNAAPRVP
              610      620      630      640      650      660

              670      680      690      700      710      720
a149-1.pep  AARLGVHLKASLTDRIDANLDYYRVFAQNKLYETRTPGHHMLNLGANYRRNTRYGEWN
|||||
m149-1      AARLGFHLKASLTDRIDANLDYYRVFAQNKLYETRTPGHHMLNLGANYRRNTRYGEWN
              670      680      690      700      710      720

              730      740      750      759
a149-1.pep  WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
|||||
m149-1      WYVKADNLLNQSVYAHSSFLSDTPQMGRSFTGGVNVKFX
              730      740      750

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 607>:

```

g150.seq (partial)
1  ..TACTGCAAGG CAGACCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT
51  CACCGCCCGC CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTGTA
101 GCGGTTTCGA TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT
151 GACAACGATC CGGCACTGGT CGGGGAAATC CTAGACCTGC TCGGCATCAA
201 TCCGGCAACG GAAATACAGG CGGGCGGAAA AACCTGCGG GTTGCTCCCG
251 CACTGTTATC CCATTCGAA CTCACGCAA ACACCCCGC CTTTGTCAA
301 GGCTATGCCA CGTTCGCCGA TAATGACGAA CTCGACCGTA TTGCTGCCGA
351 CAACGCCGTT TTGCAAGGCT TTGTGCAAAG CACGCCGATT GCCGGTGTGC
401 TGCACCGCTT CCCGGCAAAA CTGACGGCGG AACAATTCGC CGGCCTGCTG
451 CGCCCGCTTG CGCCGCGCCT GTATTGCGATT TCCTCGTCGC AGGCGGAAGC
501 GGGGGACGAA GTGCACCTGA CCGTCGGCGC AGTGCCTTTC GAACACGAAG
551 GGCGCGCCAG GCGGGGCGGC GCATCGGGTT TCTTGCCGA CCGGCTGGAA
601 GAGGACGGCA CGGTGCGCGT GTTTGCGGAA CGCAACGACG GCTTCAGGCT
651 GCCCGAAGAC AGCCGCAAGC CGATTGTGAT GATCGGCTCC GGTACCGGCG
701 TCGCACCGTT CCGCGCCTTC GTCCAACAAC GTGCCGCGA AAATGCGGAA
751 GGCAGAAACT GGCTGATTTT CGGCAATCCG CATTTTGCCG CCGACTTCCT
801 CTATCAGACC GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGAT
851 ATGACTTCGC CTGGTCGCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC
901 AAAATCCGCG AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC
951 GCATATCTAT GTGTGCGGCG ATGCGGCAAA AATGGCAAAA GAAGTGGGAA
1001 CCGCCTTGCT GGATGTGATT ATCGGGGCGG GGCATTGCGA CGAAGACGGC
1051 GCAGAAGGAT ATTTGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA
1101 TGTATTATGA

```

This corresponds to the amino acid sequence <SEQ ID 608; ORF 150.ng>:

```

g150.pep (partial)
1  ..YCKADPFPA LLANQKITAR QSDKDVRHIE IDLSGSDLHY LPGDALGVWF
51  DNDPALVGEI LDLLGINPAT EIQAGGKTLF VASALLSHFE LTQNTPAFVK
101 GYATFADNDE LDRIAADNAV LQGFVQSTPI AGVLHRFPK LTAEQFAGLL
151 RPLAPRLYSI SSSQAEAGDE VHLTVGAVRF EHEGRARAGG ASGFFADRLE
201 EDGTVRVFAE RNDGFRLPED SRKPIVMIGS GTGVAPFRAF VQRAAENAE
251 GRNWLIFGNP HFAADFLYQT EWQQFAKDFG LHRYPFAWSR DQEEKIYVQD
301 KIREQAEGLW QWLQEGAHY VCGDAAKMAK EVEAALLDVI IGAGHSDEDEG
351 AEGYLDMLRE EKRYQRDVY*

```

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 609>:

```

m150.seq
1  ATGCAGAACA CAAATCCGCC ATTACCGCCT CTGCCGCCCG AAATCACGCA
51  GCTCCTGTCT GGGCTGGACG CGGCACAATG GGCGTGGCTG TCCGGCTACG
101 CTTGGGCAAA AGCAGGAAAC GGGGCATCTG CAGGACTGCC CGCGCTTCAG
151 ACGGCATTGC CGGCGGCAGA ACCTTTTTC GTAAACCGTCC TTTCCGCCTC
201 GCAAACCGGC AATGCGAAAT CCGTTGCCGA CAAAGCGGCG GACAGCCTGG
251 AAGCCGCGCG CATCCAAGTC AGTCGCGCCG AACTGAAAGA CTATAAGGCG
301 AAAAACATCG CCGGCGAAGC CCGCCTGCTG CTGGTTACCT CCACCCAAGG
351 CGAAGGCGAA CCGCCGAAAG AAGCCGTCGT GCTGCACAAA CTGCTGAACG
401 GCAAAAAAGC CCCGAAATTG GACAACTCC AATTGCGCT ACTGGGTTTG

```


427

```

451 GGCAGCAGTT CCTATCCGAA TTTCTGTCAG GCAGGTAAG ATTCGACCG
501 GCGTTTGTAA GAATTGGGCG CAAAACGGCT GCTCGAACGC GTTGATGCGG
551 ATTTGGACTT TACCGCCTCC GAAAACGGCT GGACAGATAA TATCGCCGCA
601 CTCTTAAAAG AAGAAGCCGC AAAAAACGG GCAACGCCCG CGCCGCAGAC
651 AACGCCCCCC GCCGCGCTTC AGACGGCACC GGATGGCAGG TACTGCAAGG
701 CAGCCCCCTT TCCCGCCGCC CTGCTGGCCA ATCAGAAAAT CACCGCCCCG
751 CAATCCGATA AAGACGTGCG CCACATCGAA ATCGATTGTA GCGGTTCGGA
801 TTTGCACTAC CTCCCGGGCG ACGCGCTCGG CGTTTGGTTT GACAACGATC
851 CGGCACTGGT CAGGGAATC CTAGACCTGC TCGGCATCGA TCCGGCAACG
901 GAAATACAGG CGGGCGGAAA GATGATGCCG GTTGCGCGCG CACTTTCATC
951 TCATTTCGAA CTCACGCAA AACTCCGGC TTTCTGCAA GGCTATGCCG
1001 CGTTCGCCCC TTATGAAGAA CTCGATAAAA TCATTGCCGA TAACGCCGTT
1051 TTGCAGGATT TCGTGCAAAA CACGCCTATT GTCGATGTGC TGCACCGCTT
1101 CCCGGCAAGC CTGACGGCAG AACAATTCAT CCGTTTACTG CGTCCGCTTG
1151 CACCCCGTTT GTATTGATT TCTTCAGCAC AGGCGGAAGT GGGCGATGAA
1201 GTGCATTTAA CTGTCGGCGT GGTTCGTTT GAACACGAAG GCCGCGCCAG
1251 AACGGGCGGC GCATCGGGTT TCCTTGCCGA CCGGCTGGAA GAGGACGGCA
1301 CGGTGCGCGT GTTTGTGGAA CGCAACGACG GCTTCAGGCT GCCCGAAGAC
1351 AGCCGCAAGC CGATTGTGAT GATCGGCTCG GGCACCGCG TCGCACCGTT
1401 CCGGCTTTC GTCCAACAAC GTGCCGAGA AAATGCGGAA GGCAAAACT
1451 GGCTGATTTT CGGCAATCCG CATTTTGCCC GTGATTTTCT CTATCAAACC
1501 GAATGGCAGC AGTTTGCCAA AGACGGCTTC CTGCACAGG ACGATTTTCG
1551 CTGGTCCCGC GATCAGGAAG AAAAAATCTA TGTGCAGGAC AAAATCCGCG
1601 AACAGGCGGA AGGACTTTGG CAATGGCTGC AGGAAGGCGC GCATATCTAT
1651 GTGTGCGGCG ATGCGGCAA AATGGCAAAA GACGTGGAAG CCGCCTTGCT
1701 GGATGTGATT ATCGGGGCG GACATTTGGA CGAAGAGGCG GCAGAAGAAT
1751 ATTTGGATAT GCTGCGCGAA GAAAAACGCT ATCAGCGTGA TGTATTATGA

```

This corresponds to the amino acid sequence <SEQ ID 610; ORF 150>:

m150.pep

```

1  MQNTNPPLPP LPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ
51  TALPAAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA
101 KNIAGERLL LVTSTQGEGE PPKEAVVLHK LLNGKKAPKL DKLQFAVLGL
151 GDSSYPNFCQ AGKDFDRRFE ELGAKRLLER VDADLDFTAS ANAWTDNIAA
201 LLKEEAAKNR ATPAPQTPP AGLQAPDGR YCKAAPFPAA LLANQKITAR
251 QSDKDVRIE IDLSGSDLHY LPGAALGVWF DNDPALVREI LDLLGIDPAT
301 EIQAGGKMP VARALSSHFE LTQNTPAFVK GYAAFAHYEE LDKIIADNAV
351 LQDFVQNTPI VDLHRFPAS LTAEQFIRLL RPLAPRLYSI SSAQAEVGDE
401 VHLTVGVVRF EHEGRARTGG ASGFLADRLE EDGTVRVFVE RNDGFRLPED
451 SRKPIVMIGS GTGVAPFRAF VQRAAENAE GKNWLIIFGNP HFARDFLYQT
501 EWQQFAKDFG LHYDFAWSR DQEEKIYVQD KIREQAEGWL OWLQEGAHY
551 VCGDAAKMAK DVEAALLDVI IGAGHLEEG AEEYLDMLRE EKRYQRDVY*

```

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 150 shows 91.3% identity over a 369 aa overlap with a predicted ORF (ORF 150.ng) from *N. gonorrhoeae*:

m150/g150

	210	220	230	240	250	260
m150.pep	LLKEEAAKNRATPAPQTTPAGLQAPDGRYCKAAPFPAA	LLANQKITARQSDKDVRIE				
g150			YCKADFPFAALLANQKITARQSDKDVRIE			
			10	20	30	
	270	280	290	300	310	320
m150.pep	IDLSGSDLHYLPGAALGVWFDNDPALVREILDLLGIDPATEIQAGGKMPVARALSSHFE					
g150	IDLSGSDLHYLPGAALGVWFDNDPALVGEILDLLGINPATEIQAGGKTLFVASALLSHEE					
	40	50	60	70	80	90
	330	340	350	360	370	380
m150.pep	LTQNTPAFVKGYAAFAHYEELDKIIADNAV	LQDFVQNTPIVDVLRFPASLTAEQFIRLL				

a150.seq

1	ATGCAGAACA	CAAATCCGCC	ATTACCGCCT	ATGCCGCCCG	AAATCACGCA
51	GCTCCTGTCTG	GGGCTGGACG	CGGCACAATG	GGCGTGGCTG	TCCGGCTACG
101	CTTGGGCAAA	ACGAGGAAC	GGGGCATCTG	CAGGACTGCC	CGCGCTTCAG
151	ACGGCATTTGC	CGACGGCAGA	ACCTTTTTC	GTAACCGTCC	TTTCCGCCTC
201	GCAAACCGGC	AATGCGAAAT	CGGTTGCCGA	CAAGCGGGG	GACAGCGTGG
251	AAGCCGACATG	CATCCAAGTC	AGTCGCGCCG	AACTGAAAGA	CTATAAGGCG
301	AAAAACCATCG	CCGCGCAACG	CCGCTTGCTG	CTGTTTACCT	CCACCCAAAG
351	GAAAGCGGAA	CCGCGGSAAG	AAGCCGTCGT	GCTGCACAAA	CTGCTGAAGC
401	GCAAAAAAGC	CCCGAAATTG	GACAAACTCC	AATTTGCCGT	ACTGGGTTTG
451	GGCGACAGCT	CCTATCCGAA	TTTCTGCCGG	GCGGGCAAAG	ATTTGCACAA
501	ACGTTTTTGA	GAATTTGGCG	CAAAAACGCT	GCTCGAAGCG	GTTGATCGGG
551	ATTTGGACTT	TGCCGCGCCG	CGACAGCGAT	GGACAGATAA	TATCGCCGCA
601	CTCTTAAAG	AAGAAGCCGC	AAAAAACCGG	GCAACGCCCG	CGCCGCAGAC
651	AACGCCCCCC	GCCGCGCTTC	AGACGGCACC	GGATGGCAGG	TACTGCAAGG
701	CAGACCCCTT	TCCCGCGCTC	CTGCTGGCCA	ATCAGAAAAA	CAGCCGCCGC
751	CAATCCGATA	AAGACCTGCG	CCACATCGAA	ATCGATTTGA	GCGGTTTCGA
801	TTTGCACTAC	CTCCCGGCG	ACGCGCTCGG	CGTTTGGTTT	GACAACGATC
851	CGGCACTGGT	CAGGGAATC	CTAGACCTGC	TCGGCATCGA	TCAGGCAACG
901	GAAATACAGC	CGGCGGGAAA	AACCTTGCCG	GTTGCCCTCG	CACTGTTATC
951	CCATTTTGAA	CTCAGCGAAA	ACACCCCGCG	CTTTGTCAAA	GGCATATGCC
1001	CGTTCGCGGA	TGATGACGAA	CTCGACCGTA	TTGCTGCCGA	CACCGCGGTT
1051	TTGCAAGGCT	TTGTGCAAG	CACGCCGATT	GCCGATGTGC	TGCACCGCTT
1101	CCCGCGCAAA	CTGACAGCGT	AACAATTCG	CGGCCTACTG	CGCCCGCTTG
1151	CGCCGCGCTT	GTATTGATG	TGCTCGTCG	AGGCGGAAT	GGGGGACGAA
1201	GTGCACCTGA	CGCTCGGCG	GGTGCCTTC	GAACACGAAG	GGCGCGCCAG
1251	GGCGGGCGGC	GCATCGGGTT	TCCTTGCCGA	CCGCTTGAA	GAGGACGGCA
1301	CGGTGCGCGT	GTTTGTGAA	CGCAACGACG	GCTTCAGGCT	GCCCGAAGAC
1351	AGCCGCAAGC	CGATTGTGAT	GATCGGCTCG	GGCACCGGCG	TGCGACCGTT
1401	CCGCGGCTTC	GTCACCAAC	GTCGCGAGA	AAATGCGGAA	GGCAAAAACT
1451	GGCTGTTTTT	CGGCAATCCG	CATTTTGCCC	GTGATTTTCT	CTATCAAAAC
1501	GAATGGCAGC	AGTTTGCCAA	AGACGGCTTC	CTGCACAGAT	ACGATTTTCG
1551	CTGGTCGCGC	GATCAGGAAG	AAAAAATCTA	TGTGCGAGAC	AAAAATCCGG
1601	AACAGGCGGA	AGGACTTTGG	CAATGGCTGC	AGGAAGGCGC	GCATATCTAT
1651	GTGTGCGGCG	ATGCGGCAAA	AATGGCAAAA	GACGTGGAAG	CGCCTTTGCT

1701 GGATGTGATT ATCGGGGCAG GACATTGGA CGAAGAGGC GCAGAAGAAT
1751 ATTTGGATAT GCTGCGCGAA GAAAACGCT ATCAGCGTGA TGTTTATTGA

This corresponds to the amino acid sequence <SEQ ID 612; ORF 150.a>:

```
a150.pep
1  MQNTNPPLPP MPPEITQLLS GLDAAQWAWL SGYAWAKAGN GASAGLPALQ
51  TALPTAEPFS VTVLSASQTG NAKSVADKAA DSLEAAGIQV SRAELKDYKA
101 KNIAGERRLL LVTSTQGEGE PPEEAVVLHK LLNGKKAPKL DKLQFAVLGL
151 GDSSYPNFCR AGKDFDKRFE ELGAKRLLER VDADLDFAAA ADGWTDNIAA
201 LLKEEAAKNR ATPAPQTTTP AGLTAPDGR YCKADPFPAAL LANQKITAR
251 QSDKDVRHIE IDLSGSDLHY LPGDALGVWF DNDPALVREI LDLLGIDQAT
301 EIQAGGKTLF VASALLSHFE LTQNTPAFVK GYAPFADDE LDRIAADNAV
351 LQGFVQSTPI ADVLHRFPK LTAEQFAGLL RPLAPRLYSI SSSQAEVGDE
401 VHLTVGAVRF EHEGRARAGG ASGFLADRLE EDGTVRVFVE RNDGFRLEPED
451 SRKPIVMIGS GTGVAPFRAF VQORAAENAE GKNWLFFGNP HFARDFLYQT
501 EWQQFAKDGFLHRYDFAWSR DQEEKIYVQD KIREQAEGLW QWLQEGAHYI
551 VCGDAAKMAK DVEAALLDVI IGAGHLDEEG AEEYLDMLRE EKRYQRDVY*
```

ml150/a150 94.8% identity in 599 aa overlap

	10	20	30	40	50	60
m150.pep	MQNTNPPLPPPLPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQALPAAEPFS					
a150	MQNTNPPLPPPLPPEITQLLSGLDAAQWAWLSGYAWAKAGNGASAGLPALQALPAAEPFS					
	10	20	30	40	50	60
m150.pep	70	80	90	100	110	120
a150	70	80	90	100	110	120
m150.pep	VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLVLTSTQGEGE					
a150	VTVLSASQTGNAKSVADKAADSLEAAGIQVSRAELKDYKAKNIAGERRLLVLTSTQGEGE					
	70	80	90	100	110	120
m150.pep	130	140	150	160	170	180
a150	130	140	150	160	170	180
m150.pep	PPKEAVVLHKLNGKKAPKLDKLQFAVLGLDSSYPNFCRAGKDFDRFEELGAKRLLER					
a150	PPEEAVVLHKLNGKKAPKLDKLQFAVLGLDSSYPNFCRAGKDFDKRFEELGAKRLLER					
	130	140	150	160	170	180
m150.pep	190	200	210	220	230	240
a150	190	200	210	220	230	240
m150.pep	VDADLDFASANAATDNIAALLKEEAAKNRATPAPQTTTPAGLTAPDGRYCKAAPPFAA					
a150	VDADLDFAAAADGWTDNIAALLKEEAAKNRATPAPQTTTPAGLTAPDGRYCKADPFPA					
	190	200	210	220	230	240
m150.pep	250	260	270	280	290	300
a150	250	260	270	280	290	300
m150.pep	LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFNDNDPALVREILDLLGIDPAT					
a150	LLANQKITARQSDKDVRHIEIDLSGSDLHYLPGDALGVWFNDNDPALVREILDLLGIDQAT					
	250	260	270	280	290	300
m150.pep	310	320	330	340	350	360
a150	310	320	330	340	350	360
m150.pep	EIQAGGKMPVARALSSHFEITQNTPAFVKGYAAFAHYEELDKIADNAVLDQFVQNTPI					
a150	EIQAGGKTLFVASALLSHFEITQNTPAFVKGYAPFADDELDRIAADNAVLDQGFVQSTPI					
	310	320	330	340	350	360
m150.pep	370	380	390	400	410	420
a150	370	380	390	400	410	420
m150.pep	VDVLRFPASLTAEQFIRLLRPLAPRLYSISSQAEVGDEVHLLTVGVVRFEEHGRARTGG					
a150	ADVLRFPKLTAEQFAGLLRPLAPRLYSISSQAEVGDEVHLLTVGAVRFEHGRARAGG					
	370	380	390	400	410	420
m150.pep	430	440	450	460	470	480
a150	430	440	450	460	470	480
m150.pep	ASGFLADRLEEDGTVRVFVERNDGFRLEPDSRKPIVMIGSGTGVPFRAFVQORAAENAE					
a150	ASGFLADRLEEDGTVRVFVERNDGFRLEPDSRKPIVMIGSGTGVPFRAFVQORAAENAE					
	430	440	450	460	470	480
m150.pep	490	500	510	520	530	540
a150	GKNWLIFGNPHFARDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGLW					

430

```

a150      GKNWLFPGNPHFARDFLYQTEWQQFAKDGFLHRYDFAWSRDQEEKIYVQDKIREQAEGW
           490       500       510       520       530       540

m150.pep  QWLQEGAHYVCGDAAKMAKDVEAALLDVIIGAGHLDEEGAEYLDMLREEKRYQRDVYX
           550       560       570       580       590       600
|||||
a150      QWLQEGAHYVCGDAAKMAKDVEAALLDVIIGAGHLDEEGAEYLDMLREEKRYQRDVYX
           550       560       570       580       590       600

```

The following partial DNA sequence was identified in *N.gonorrhoeae* <SEQ ID 613>:

```

g151.seq
1  ATGAAACAAA TCCGCAACAT CGCCATCATC GCACACGTCG ACCACGGCAA
51  AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGA AAAAGAA
151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTG
201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GCGCGCGAAG
251 TGGAGCGCGT TTTGGGGATG GTGGATTGCG TCGTCTTGTT GGTGGACGCA
301 CAGGAAGGTC CGATGCCGCA AATCCGTTTC GTGACCAAAA AAGCCTTGGC
351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAACCGTCCG
401 CCCGTCCGAG CTGGGTTATC GACCAGACTT TCGAGTTGTT CGACAACCTG
451 GGTGCGACCG ACGAGCAGTT GGATTTCCTG ATTGTTTACG CTTCAGGTTT
501 GACGCGCTTT GCCAAGCTGG AAGAAAccga CGAGAGCAGC GATATGCGCC
551 CGCtggttcgA CACCATCCTA AAATACacgc ctgCACCAG CGGCAGCGCG
601 GACGAGCCGC TGCAACTGCA AATTTCCCAA CTCGACTACG ACAACTACAC
651 CGGCCGCCCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCGGCC
701 AAACCGTTGC CGTGATGAAC CACGAGCAGC AAATCGCCCA AGGCCGCATC
751 AACCAGCTTT TGGGTTTCAA AGGCTTGGA CGCGTGCCGC TTGAAGAAGC
801 CGAAGCCGGC GACATTGTGA TTATTTCCGG TATCGAAGAC ATCGGCATCG
851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC
901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTAAACA CCAGCCCGCT
951 CGCAGGTACA GAAGGCAAAT TCGTGACCAG CCGCCAAATC CGCGACCGCC
1001 TGCAAAAAGA ATGTGTGACC AACGTTGCCC TGCGCGTGGA AGACACCGCC
1051 GatgCGGACG TGTTCCGCGT ATCcgGGCGC GCGCAACTGC ACCTGACGAT
1101 TTTGCTGGAA AATATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGAAGC
1151 CGCGCGTCGT GTACCGAGAC ATCGACGGTC AAAAAAGCGA ACCTTATGAA
1201 AACCTGACTG TGGACGTACC cgacgacAAC CAAGGCGCGG TAATGGAAGA
1251 ACTCGGCGCG CGCCGTGGCG AACTGACCAA TATGGAAAGC GACGGAACG
1301 GacgCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTT
1351 CAAGGCGAAT TCATGACCCT GACGCGCGGC GTCGGGCTGA TGAgccacGT
1401 GTTcgacgac tacgcgcccg tcaAACCCGA TATGCCCGGC CGCCACAACG
1451 GCGTactggt GtcccaAGAG CAGGGCGAGG CGGTGCTTA CGCCTTGTGG
1501 AATCTTGAAG ACCGCGGCCG TATGTTGTA TCGCCCAACG ACAAATCTA
1551 CGAAGGTATG ATTATCGGCA TCCACAGCCG CGACAACGAT TTGGTGGTCA
1601 ACCCGCTCAA AGGCAAAAA CTCACCAATA TCCGTGCCAG CGGTACCGAC
1651 GAAGCGGTGC GCCTGACCAC GCCGATCAA CTGAcgtTGG AAGGCGCGGT
1701 CGAGTTTATC GACGATGACG AGCTGGTGGA AATCACGCCG CAAtccatcc
1751 gcctgcgcat gcgttacctG AGCGaattgg aacgcccgcg tcaTTTAA
1801 AagctgGATT AA

```

This corresponds to the amino acid sequence <SEQ ID 614; ORF 151.ng>:

```

g151.pep
1  MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
51  RGITILAKNT AIDYEGCHIN IVDTPGHADF GGEVERVLGM VDCVLLVDA
101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
151 GATDEQLDFP IVYASGLSGF AKLEETDESS DMRPLFDITL KYTPAPSGSA
201 DEPLQLQISQ LDYDNYTGRL GIGRIILNGRI KPGQTVAVMN HEQQIAQGRI
251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
401 NLTVDPDDN QGAVMEELGR RRGELTNMES DGNGRTRLEY HIPARGLIGF
451 QGEFMTLTRG VGLMSHVDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW

```

501 NLEDGRGMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRMYL SELERRRHF
 601 KLD*

The following partial DNA sequence was identified in *N.meningitidis* <SEQ ID 615>:

m151.seq
 1 ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA
 51 AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
 101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGA AAAAGAA
 151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
 201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGGCGAAG
 251 TAGAGCGCGT TTTGGGGATG GTGGACTGCG TCGTCTTGTT GGTGGACGCG
 301 CAGGAAGGCC CGATGCCGCA AACCCGTTTC GTGACCAAAA AAGCCTTGGC
 351 TTTGGGGCTG AAACCGATTG TCGTCATCAA CAAAATCGAC AAGCCGTCCG
 401 CTCGTCGAG CTGGGTTATC GACCAAACTT TCGAGCTGTT CGACAACCTG
 451 GGCGCGACCG ACGAGCAGTT GGATTCCCG ATTGTTTACG CTCAGGGTT
 501 GAGCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
 551 CGCTGTTTCA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
 601 GACGAAACGC TGCAACTGCA AATTTCCTCA CTGCACTACG ACAACTACAC
 651 CGGCCCGCTC GGTATCGGTC GTATCTTGAA CGGACGCATC AAACCCGGCC
 701 AAACCGTTGC CGTCATGAAC CACGATCAGC AAATCGCCCA AGGCCGCATC
 751 AACCAGTTTT TGGGTTTCAA AGGTTTGAA CGCGTGCCGC TTGAAGAAGC
 801 CGAAGCCGGC GACATCGTGA TTATTTCCGG TATCGAAGAC ATCGGTATCG
 851 GCGTAACCAT CACCGACAAA GACAATCCCA AAGGCCTACC GATGTTGAGC
 901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGCT
 951 GGCGGGTACG GAAGGCAAT TCGTAACCA CCGCAAAATC CGCGACCGCC
 1001 TGCAAAAAGA ATTGCTGACC AACGTCGCC TGCGCGTGGA AGATACCGCC
 1051 GATGCCGACG TGTTCCGCGT ATCCGGGCGC GCGGAGCTGC ACCTGACCAT
 1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAAC
 1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
 1201 AACCTGACCG TGGATGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
 1251 ACTCGGCCGC CGCCGTGGCG AACTGACTAA TATGGAAGC GACGGCAACG
 1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGTTTC
 1351 CAAGGCGAAT TTATGACCTT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
 1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCCCGGC CGCCACAACG
 1451 GCGTGCTGGT GTCCCAAGAG CAGGCGGAGG CAGTCGCTTA CGCCTTGTTG
 1501 AATCTGGAAG ACCGCGGCCG TATGTTCTGA TCGCCCAACG ACAAATCTA
 1551 CGAAGGCATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
 1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
 1651 GAAGCGTTTC GCCTGACCAC GCCAATCAAG CTGACGCTGG AAGGTGCGGT
 1701 TGAGTTTATC GACGATGACG AACTCGTTGA AATCACGCCG CAATCCATCC
 1751 GTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCCG CCACTTTAAA
 1801 AAGCTGGATT GA

This corresponds to the amino acid sequence <SEQ ID 616; ORF 151>:

m151.pep
 1 MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
 51 RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVLLLVDA
 101 QEGPMPQTRF VTKKALALGL KPIVINKID KPSARPSWVI DQTFELFDNL
 151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDITL KYTPAPSGSA
 201 DETLQLQISQ LDYDNYTGRL GIGRIINGRI KPGQTVAVMN HDQQIAQGRI
 251 NQLLGFKGLE RVPLEAEAG DIVIISGIED IGIGVTITDK DNPKGLPMLS
 301 VDEPTLTMDF MVNTSPLAGT EGKFVTSRQI RDRLQKELT NVALRVEDTA
 351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
 401 NLTVDVPDDN QGAVMEELGR RRGELTNMES DNGRTRLEY HIPARGLIGF
 451 QGEFMTLTRG VGLMSHVDDD YAPVKPDMPG RHNGVLVSQE QGEAVAYALW
 501 NLEDGRGMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGKK LTNIRASGTD
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRKRYL SELERRRHF
 601 KLD*

Computer analysis of this amino acid sequence gave the following results:

Homology with a predicted ORF from *N.gonorrhoeae*

ORF 151 shows 99.2% identity over a 603 aa overlap with a predicted ORF (ORF 151.ng)
from *N. gonorrhoeae*:

m151/g151

m151.pep	10	20	30	40	50	60
	MKQIRNIAIIAHVDHGKTTLV	DQLLRQSGTFRANQQVDERV	MDSNDLEKERGITILAKNT			
g151	MKQIRNIAIIAHVDHGKTTLV	DQLLRQSGTFRANQQVDERV	MDSNDLEKERGITILAKNT			
m151.pep	70	80	90	100	110	120
	AIDYEGYHINIVDTPGHADFG	GEVERVLGMVDCVLLVDAQ	EGMPQTRFVTKKALALGL			
g151	AIDYEGCHINIVDTPGHADFG	GEVERVLGMVDCVLLVDAQ	EGMPQTRFVTKKALALGL			
m151.pep	130	140	150	160	170	180
	KPIVVINKIDKPSARPSWVID	QTFELFDNLGATDEQLDFPI	VYASGLSGFAKLEETDESN			
g151	KPIVVINKIDKPSARPSWVID	QTFELFDNLGATDEQLDFPI	VYASGLSGFAKLEETDESS			
m151.pep	190	200	210	220	230	240
	DMRPLFDITILKYTPAPSGS	ADELQLQISQLDYDNYTGRL	GIGRILNGRIKPGQTVAVMN			
g151	DMRPLFDITILKYTPAPSGS	ADELQLQISQLDYDNYTGRL	GIGRILNGRIKPGQTVAVMN			
m151.pep	250	260	270	280	290	300
	HDQQIAQGRINQLLGFKGL	ERVPLEEAEAGDIVIISGIE	DIGIGVTITDKDNPKGLPMLS			
g151	HEQQIAQGRINQLLGFKGL	ERVPLEEAEAGDIVIISGIE	DIGIGVTITDKDNPKGLPMLS			
m151.pep	310	320	330	340	350	360
	VDEPTLTMDFMVNTSPLAG	TEGKFVTSRQIRDLQKELL	TNVALRVEDTADADVFRVSGR			
g151	VDEPTLTMDFMVNTSPLAG	TEGKFVTSRQIRDLQKELL	TNVALRVEDTADADVFRVSGR			
m151.pep	370	380	390	400	410	420
	GELHLTILLENMRREGYEL	AVGKPRVVYRDIDGQKCEPY	ENLTVDPDDNQGAVMEELGR			
g151	GELHLTILLENMRREGYEL	AVGKPRVVYRDIDGQKCEPY	ENLTVDPDDNQGAVMEELGR			
m151.pep	430	440	450	460	470	480
	RRGELTNMESDGNRTRLEY	HIPARGLIGFQGEFMTL	TRGVGLMSHVDDYAPVKP	DPMPG		
g151	RRGELTNMESDGNRTRLEY	HIPARGLIGFQGEFMTL	TRGVGLMSHVDDYAPVKP	DPMPG		
m151.pep	490	500	510	520	530	540
	RHNGVLVSQEQGEAVAYAL	WNLEDRGMFVSPNDKIYEG	MIIGIHSRDNDLVVNPLK	GKK		
g151	RHNGVLVSQEQGEAVAYAL	WNLEDRGMFVSPNDKIYEG	MIIGIHSRDNDLVVNPLK	GKK		
m151.pep	550	560	570	580	590	600
	LTNIRASGTDEAVRLTTP	IKLTLEGAVEFIDDELVEIT	PQSIRLRKRYLSELERRRH	FK		
g151	LTNIRASGTDEAVRLTTP	IKLTLEGAVEFIDDELVEIT	PQSIRLRMYLSELERRRH	FK		

m151.pep KLDX
 ||||
 g151 KLDX

The following partial DNA sequence was identified in *N. meningitidis* <SEQ ID 617>:

a151.seq
 1 ATGAAACAAA TCCGCAACAT CGCCATCATC GCCACGTCG ACCACGGCAA
 51 AACCACATTG GTCGACCAAC TGCTGCGCCA ATCCGGCACA TTCCGCGCCA
 101 ACCAGCAGGT TGACGAGCGC GTGATGGACA GCAACGACCT TGAAGAAGAA
 151 CGCGGCATCA CCATCCTCGC CAAAAACACC GCCATCGATT ACGAAGGCTA
 201 CCACATCAAT ATCGTCGACA CGCCGGGACA CGCCGACTTC GGCGCGAAG
 251 TAGAGCGAGT TTTGGGGATG GTGGACTGCG TCGTCTTGT GGTGGACGCG
 301 CAGGAAGGCC CGATGCCGCA AACCCTTTC GTGACCAAAA AAGCCTTGCC
 351 TTTGGGGCTG AAACCGATTG TCGTCATCAA TAAATCGAC AAACCTCCG
 401 CCGTCCGAG CTGGGTCATC GACCAACTT TCGAGCTGTT CGACAACTTG
 451 GGCGCGACCG ACGAGCAGTT GGATTCCCG ATTGTTTATG CTTCCGGTCT
 501 GTCCGGTTTC GCCAAATTGG AAGAAACCGA CGAGAGCAAC GACATGCGTC
 551 CGCTGTTCGA TACTATCTTA AAATATACGC CTGCACCGAG CGGCAGCGCG
 601 GACGAAACGC TGCAACTGCA AATTTCCTAA CTCGACTACG ACAACTACAC
 651 CGGCCGCCTC GGTATCGGTC GTATCTTGAA CGGACGTATC AAGCCCGGTC
 701 AAGTTGTGTC CGTCATGAAC CACGATCAAC AAATCGCCCA AGGCCGCATC
 751 AACCAGCTTT TGGGTTTCAA AGGTTAGAA CGCGTGCCCG TTGAAGAAGC
 801 CGAAGCCGGC GACATCGTGA TTATTCCGG TATTGAAGAC ATCGGCATCG
 851 GCGTAACCAT CACCGACAAA GACAACCCCA AAGGCCTGCC GATGTTGAGC
 901 GTGGACGAAC CGACGCTGAC GATGGACTTT ATGGTCAACA CCAGCCCGTT
 951 GGCAGGTACG GAAGGCAAAT TCGTAACCA GCGCCAAATC CGCGACCGCC
 1001 TGCAAAAAGA ATTGCTGACC AACGTCGCCC TGCGCGTGGA AGATACCGCC
 1051 GATGCCGACG TGTTCCGCGT ATCCGGGCGC GGCGAGCTGC ACCTGACCAT
 1101 TTTGCTGGAA AACATGCGCC GCGAAGGCTA CGAACTCGCC GTCGGCAAAC
 1151 CGCGCGTCGT GTACCGCGAC ATCGACGGTC AAAAATGCGA ACCGTATGAA
 1201 AACCTGACCG TGGACGTACC CGACGACAAC CAAGGCGCGG TAATGGAAGA
 1251 ACTCGGCCCG CGCCGTGGCG AACTGACTAA TATGGAAAGC GACGGCAACG
 1301 GACGCACCCG CCTCGAATAC CATATTCCAG CGCGCGGCTT GATCGGCTTC
 1351 CAAGGCGAAT TTATGACCCT GACGCGCGGG GTCGGGCTGA TGAGCCACGT
 1401 GTTCGACGAT TACGCGCCCG TCAAACCCGA TATGCTTGGC CGCCACAACG
 1451 GCGTGCTGGT GTCCCAAGAG CAGGGCGAGG CAGTCGCTTA CGCCTTGTGG
 1501 AATCTGGAAG ACCGCGGCCG TATGTTCTGA TCGCCCAACG ACAAATCTA
 1551 CAAGGTATG ATTATCGGCA TCCACAGTCG CGACAACGAT TTGGTGGTCA
 1601 ACCCGCTCAA AGGCAAAAAA CTTACCAACA TCCGTGCCAG CGGTACCGAC
 1651 GAAGCCGTTT GCCTGACCAC GCCGATTAAG CTGACGCTGG AAGGTGCGGT
 1701 CGAGTTTATC GACGATGATG AGCTGGTAGA AATCACGCCG CAATCCATCC
 1751 CTCTGCGCAA GCGTTACTTG AGCGAATTGG AACGCCGCGC CCATTTCAAA
 1801 AAGCTAGATT GA

This corresponds to the amino acid sequence <SEQ ID 618; ORF 151.a>:

a151.pep
 1 MKQIRNIAII AHVDHGKTTL VDQLLRQSGT FRANQQVDER VMDSNDLEKE
 51 RGITILAKNT AIDYEGYHIN IVDTPGHADF GGEVERVLGM VDCVLLVDA
 101 QEGPMPQTRF VTKKALALGL KPIVVINKID KPSARPSWVI DQTFELFDNL
 151 GATDEQLDFP IVYASGLSGF AKLEETDESN DMRPLFDLIL KYTPAPSGSA
 201 DETLQLQISQ LDYDNYTGRL GIGRILNGRI KPGQVVAVMN HDQQAQGRRI
 251 NQLLGFKGLE RVPLEEAEAG DIVIISGIED IGIGVTITDK DNPGLPLMLS
 301 VDEPTLTMDL MVNTSPLAGT EGKFVTSRQI RDRLQKELLT NVALRVEDTA
 351 DADVFRVSGR GELHLTILLE NMRREGYELA VGKPRVVYRD IDGQKCEPYE
 401 NLTVDVDDN QGAVMEELGR RRGELTNMES DNGRTRLEY HIPARGLIGF
 451 QGEFMTLTRG VGLMSHVFDL YAPVKPDMFG RHNGVLVSQE QGEAVAYALW
 501 NLEDRGRMFV SPNDKIYEGM IIGIHSRDND LVVNPLKGGK LTNIRASGTD
 551 EAVRLTTPIK LTLEGAVEFI DDELVEITP QSIRLRKRYL SELERRRHFK
 601 KLD*

m151/a151 99.8% identity in 603 aa overlap

10 20 30 40 50 60